```
              GenCore version 5.1.6
        Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      October 27, 2005, 15:26:51 ; Search time 274 Seconds
                                           (without alignments)
                                           4042.634 Million cell updates/sec

Title:          US-10-009-002-2
Perfect score:  15376
Sequence:     1 MPVISTQTSPVPAPRTRKNK.........KYLAVIVFALGLIAVGLAIS 2864

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:     2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters:      2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    A_Geneseq_16Dec04:*
              1: geneseqp1980s:*
              2: geneseqp1990s:*
              3: geneseqp2000s:*
              4: geneseqp2001s:*
              5: geneseqp2002s:*
              6: geneseqp2003a8:*
              7: geneseqp2003b8:*
              8: geneseqp2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	15376	100.0	2864	4	AAB59171	Aab59171 Protein e
2	15366	99.9	2864	6	ADA77730	Ada77730 Genomic h
3	15317	99.6	2864	2	AAR82072	Aar82072 Hepatitis
4	15317	99.6	2864	3	AAB09268	Aab09268 Hepatitis
5	7380	48.0	1422	3	AAB09038	Aab09038 Hepatitis
6	7375	48.0	1422	2	AAR82068	Aar82068 Hepatitis
7	3277.5	21.3	3033	4	AAB59172	Aab59172 Protein e
8	3277.5	21.3	3033	4	AAB31168	Aab31168 Amino aci
9	3277.5	21.3	3033	3	AAB30729	Aab30729 Amino aci
10	3268.5	21.3	3033	2	AAR33539	Aar33539 NANBH vir
11	3265.5	21.2	3033	5	ABG30688	Abg30688 Human HCV
12	3263	21.2	3010	5	ABG32457	Abg32457 Hepatitis
13	3262	21.2	3010	5	ABG32451	Abg32451 Hepatitis
14	3262	21.2	3010	5	ABG32460	Abg32460 Hepatitis
15	3262	21.2	3010	5	ABG32453	Abg32453 Hepatitis
16	3262	21.2	3010	5	ABG32461	Abg32461 Hepatitis
17	3260	21.2	3010	5	ABG32458	Abg32458 Hepatitis
18	3260	21.2	3010	5	ABG32459	Abg32459 Hepatitis
19	3260	21.2	3033	2	AAR33538	Aar33538 NANBH vir
20	3258	21.2	3010	5	ABG32454	Abg32454 Hepatitis
21	3256	21.2	3010	5	ABG32455	Abg32455 Hepatitis
22	3256	21.2	3010	5	ABG32452	Abg32452 Hepatitis
23	3256	21.2	3010	4	AAB59174	Aab59174 Protein e
24	3252.5	21.2	3011	2	ABG32456	Abg32456 Hepatitis
25	3247	21.1	3010	2	AAW98022	Aaw98022 Infectiou
26	3247	21.1	3010	4	AAB31170	Aab31170 Amino aci
27	3247	21.1	3010	8	ADO36227	Ado36227 Hepatitis
28	3247	21.1	3010	8	ADO79401	Ado79401 Hepatitis
29	3234.5	21.0	3011	5	AAU84597	Aau84597 HCV polyp
30	3232	21.0	3010	5	AAE20477	Aae20477 HCV-S1 fu
31	3228	21.0	2984	4	AAE00447	Aae00447 Hepatitis
32	3226	21.0	2984	4	AAE00442	Aae00442 Hepatitis
33	3223	21.0	3010	7	AAY06423	Aay06423 Non-A, no
34	3223	21.0	3010	7	ADF88597	Adf88597 Hepatitis
35	3222.5	21.0	3011	2	AAR66995	Aar66995 Hepatitis
36	3220	20.9	2984	4	AAE00449	Aae00449 Hepatitis
37	3218	20.9	3010	2	AAR20111	Aar20111 Non-A, no
38	3218	20.9	3010	2	AAR20091	Aar20091 Non-A, no
39	3217	20.9	3010	2	AAR68864	Aar68864 Hepatitis
40	3216	20.9	3010	2	AAR68622	Aar68622 HCV prote
41	3216	20.9	3010	2	AAR82694	Aar82694 Partial H
42	3214.5	20.9	3011	4	AAB59173	Aab59173 Protein e
43	3214.5	20.9	3011	4	AAB31169	Aab31169 Amino aci
44	3214.5	20.9	3011	5	AAU79221	Aau79221 Hepatitis
45	3214.5	20.9	3011	8	ADJ56744	Adj56744 Hepatitis

ALIGNMENTS

```
RESULT 1
AAB59171
ID   AAB59171 standard; protein; 2864 AA.
XX
AC   AAB59171;
XX
DT   11-SEP-2003 (revised)
DT   21-MAR-2001 (first entry)
XX
DE   Protein encoded by infectious GB virus B genome.
XX
KW   GBV-B; hepatitis C virus; HCV; vaccine.
XX
OS   Hepatitis GB virus B.
XX
PN   WO200075337-A1.
XX
PD   14-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015293.
XX
PR   04-JUN-1999; 99US-0137694P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Bukh J, Yanagi M, Emerson SU, Purcell RH;
XX
DR   WPI; 2001-091214/10.
XX
PT   New infectious nucleic acids of the GB virus-B clone, useful for
PT   indirectly studying the molecular properties of hepatitis C virus (HCV)
PT   and in developing vaccines and therapeutics for HCV.
XX
PS   Disclosure; Page 64-75; 96pp; English.
XX
CC   The present invention relates to GB virus-B. The nucleic acid molecules
CC   of the invention are useful for indirectly studying the molecular
CC   properties of hepatitis C virus (HCV). The infectious nucleic acid
CC   sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used
CC   in the development of vaccines and therapeutics for HCV. (Updated on 11-
CC   SEP-2003 to standardise OS field)
XX
SQ   Sequence 2864 AA;

Query Match             100.0%;  Score 15376;  DB 4;  Length 2864;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 2864;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy      1  MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD   60

Qy     61  GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPV   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPV   120

Qy    121  CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY  180

Qy    181  CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG  240

Qy    241  ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK  300

Qy    301  LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI  360

Qy    361  AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF  420

Qy    421  KNNTWGGCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAIIQYPG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  KNNTWGGCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAIIQYPG  480

Qy    481  SKEMFKPHNWMSGHLYPEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  SKEMFKPHNWMSGHLYPEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ  540

Qy    541  GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY  600

Qy    601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL  660

Qy    661  LGFVPMAAGLPLTFFVAAAAQPDYDWMWVRLLVAGLVLVWAGRNRGHRIALLVGPWPLVAL  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  LGFVPMAAGLPLTFFVAAAAQPDYDWMWVRLLVAGLVLVWAGRNRGHRIALLVGPWPLVAL  720

Qy    721  LTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLMQRWENWF  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  LTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLMQRWENWF  780

Qy    781  WNVTLRPERFFLVLVCFPGATYDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  WNVTLRPERFFLVLVCFPGATYDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR  840

Qy    841  LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPEGKARVY   900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPEGKARVY   900

Qy    901  RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPPTLQCLSERGTLSAMAVV  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPPTLQCLSERGTLSAMAVV  960

Qy    961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020

Qy   1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDDYWCVCCGALPMAVAKGSSG 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDDYWCVCCGALPMAVAKGSSG 1080

Qy   1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ 1140

Qy   1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200

Qy   1201  TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT 1260

Qy   1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKGHCDELANELA 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKGHCDELANELA 1320

Qy   1321  RKGITAVSYYRGCDISKIPEGDCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  RKGITAVSYYRGCDISKIPEGDCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT  1380

Qy   1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYVVDGSCTPSGMVPECNIVEAFDAAKAWYG 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYVVDGSCTPSGMVPECNIVEAFDAAKAWYG 1440

Qy   1441  LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNVVLLTAAQ 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441  LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNVVLLTAAQ 1500

Qy   1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT 1560

Qy   1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEECASFIP  1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEECASFIP  1620

Qy   1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILALIEYCCGLVTLPDNPFA 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILALIEYCCGLVTLPDNPFA 1680

Qy   1681  SCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAGTALGTWTSVGF  1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681  SCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAGTALGTWTSVGF  1740

Qy   1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT 1800

Qy   1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE 1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE 1860

Qy   1861  DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861  DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ 1920

Qy   1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVNYG 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVNYG 1980

Qy   1981  VRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVADGVQVQCYLGEPKTPWTTSAC  2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981  VRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVADGVQVQCYLGEPKTPWTTSAC  2040

Qy   2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE 2100

Qy   2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL 2160

Qy   2161  QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK 2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
XX
PR   15-JAN-2002; 2002US-0348573P.
PR   06-JUN-2002; 2002US-0386655P.
XX
PA   (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
PI   De Tomassi A, Graziani R, Paonessa G, Traboni C;
XX
DR   WPI; 2003-598503/56.
DR   N-PSDB; ADA77752.
XX
PT   New GB virus B (GBV-B) replicon for identifying compounds that inhibit
PT   GBV-B or hepatitis C virus, comprises a GBV-B 5' UTR, a selection or
PT   reporter sequence, an internal ribosome entry site, an NS3-NSSB sequence,
PT   and a GBV-B 3' UTR.
XX
PS   Disclosure; Fig 3; 81pp; English.
XX
CC   This invention relates to a novel GB virus-B (GBV-B) replicon and
CC   replicon enhanced cells. A GBV-B replicon is an RNA molecule able to
CC   autonomously replicate in a cultured cell to produce detectable levels of
CC   one or more GBV-B proteins. Specifically, it may comprise the GBV-B 5'
CC   UTR, GBV-B structural region, selection or reporter sequence, internal
CC   ribosome entry site, NS3-NSSB sequence, and GBV-B 3' UTR. Accordingly,
CC   they are useful in providing tools for studying GBV-B replication,
CC   polyprotein production and processing, identifying compounds that inhibit
CC   GBV-B, providing a surrogate model for identifying compounds that inhibit
CC   HCV, and providing a scaffold for producing GBV-B/HCV chimeric replicons.
CC   Due to the similarity between GBV-B and the hepatitis C virus (HCV),
CC   compounds that inhibit GBV-B may be useful antiviral agents, specifically
CC   anti-HCV agents. The GBV-B subgenomic replicon constructs termed GBV-B-
CC   neo-RepA (neo-RepA), neo-RepB, neo-RepC and neo-RepD were produced by
CC   replacing the regions coding for structural proteins and the NS2 protein
CC   with the sequences of neomycin phosphotransferase gene (neo) and
CC   encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) in
CC   the plasmid FL3/pACYC177. This polypeptide sequence is the GBV-B replicon
CC   protein sequence of the invention.
XX
SQ   Sequence 2864 AA;

Query Match             99.9%;  Score 15366;  DB 6;  Length 2864;
Best Local Similarity   99.9%;  Pred. No. 0;
Matches 2860;  Conservative   4;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1  MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRVQRDARPRNYKIAGIHD   60
         ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    1  MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD   60

Qy   61  GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLVGPLVAGAVVRPV              120
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLVGPLVAGAVVRPV              120

Qy  121  CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY  180

Qy  181  CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG  240

Qy  241  ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK  300

Qy  301  LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI  360

Qy  361  AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF  420
         |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db  361  AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPVTLEYNNSISWYPYTIPGARGCMVKF  420

Qy  421  KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG  480
```

```
Db  2161  QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK  2220
Qy  2221  EVSEWSDESWSTATTASSYVTGPPYPYKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  EVSEWSDESWSTATTASSYVTGPPYPYKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280

Qy  2281  TDVISFKTASKVLSATRAITSGFLKQRSLVVTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TDVISFKTASKVLSATRAITSGFLKQRSLVVTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340

Qy  2341  VRLAKEKASKVGVMWDYDEVAAHTPSKSAAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  VRLAKEKASKVGVMWDYDEVAAHTPSKSAAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400

Qy  2401  GEIPSHYRQTVIVPKBEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV  2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  GEIPSHYRQTVIVPKBEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV  2460

Qy  2461  MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520

Qy  2521  RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAABQAGMK  2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAABQAGMK  2580

Qy  2581  NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQKYSLEELTSCSSN  2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQKYSLEELTSCSSN  2640

Qy  2641  VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHYPCLWVSRVLAVHF  2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHYPCLWVSRVLAVHF  2700

Qy  2701  MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL  2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL  2760

Qy  2761  TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN  2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN  2820

Qy  2821  YCDVYSPEGDVFITPQRRLQKFLVKYLAVIVPALGLIAVGLAIS  2864
          ||||||||||||||||||||||||||||||||||||||||||||
Db  2821  YCDVYSPEGDVFITPQRRLQKFLVKYLAVIVPALGLIAVGLAIS  2864
```

```
RESULT 2
ADA77730
ID   ADA77730 standard; protein; 2864 AA.
XX
AC   ADA77730;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Genomic hepatitis GB virus B (GBV-B) replicon protein sequence.
XX
KW   GB virus-B; GBV-B; replicon; autonomous replication; NS3-NS5B; HCV;
KW   hepatitis C virus; antiviral; anti-HCV; neomycin phosphotransferase gene;
KW   neo.
XX
OS   Hepatitis GB virus B.
XX
FH   Key             Location/Qualifiers
FT   Protein         2275..2864
FT                   /note= "Non structural protein 5B, RNA dependent RNA
FT                   polymerase"
XX
PN   WO200305944-A2.
XX
PD   24-JUL-2003.
XX
PF   13-JAN-2003; 2003WO-EP000281.
```

```
Qy    421  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  480
Db    421  KNNTWGCCRIRRNVPSYCTMGGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG  480

Qy    481  SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ  540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ  540

Qy    541  GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY  600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY  600

Qy    601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL  660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL  660

Qy    661  LGFVPMAAGLPLTFFVAAAAQPDYDWMWVRLLVAGLVLWAGRNRGHRIALLVGPWPLVAL  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  LGFVPMAAGLPLTFFVAAAAQPDYDWMWVRLLVAGLVLWAGRDRGHRIALLVGPWPLVAL  720

Qy    721  LTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLMQRWENWF  780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  LTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLMQRWENWF  780

Qy    781  WNVTLRPERFFLVLVCFPGATYDALVTFCVCHVALLCLTSSAAASFFGTDSRVRAHRMLVR  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  WNVTLRPERFFLVLVCFPGATYDALVTFCVCHVALLCLTSSAAASFFGTDSRVRAHRMLVR  840

Qy    841  LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPPEGKARVY  900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPPEGKARVY  900

Qy    901  RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPFTLQCLSERGTLSAMAVV  960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPFTLQCLSERGTLSAMAVV  960

Qy    961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPGSIHPITVDAAN  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPGSIHPITVDAAN  1020

Qy   1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG  1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG  1080

Qy   1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ  1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ  1140

Qy   1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN  1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN  1200

Qy   1201  TGASLITYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT  1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  TGASLITYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT  1260

Qy   1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKLKKGRHLIFEATKKGHCDELANELA  1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKLKKGRHLIFEATKKGHCDELANELA  1320

Qy   1321  RKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT  1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  RKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT  1380

Qy   1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCCTPSGMVPECNIVEAFDAAAKWYG  1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCCTPSGMVPECNIVEAFDAAAKWYG  1440

Qy   1441  LSSTEAQFILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTTAAQ  1500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441  LSSTEAQFILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTTAAQ  1500

Qy   1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT  1560
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT  1560
```

```
Db   1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT  1560

Qy   1561  SGTAALAVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEECASFIP  1620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561  SGTAALAVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEECASFIP  1620

Qy   1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA  1680
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA  1680

Qy   1681  SCVFAFIAGITTTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF  1740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681  SCVFAFIAGITTTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF  1740

Qy   1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT  1800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT  1800

Qy   1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEEASTPWSVISACIRWLHTPTE  1860
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEEASTPWSVISACIRWLHTPTE  1860

Qy   1861  DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ  1920
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861  DDCCLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ  1920

Qy   1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVNYG  1980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVNYG  1980

Qy   1981  VRDCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVADGVVQVCYLGEPKTPWTTSAC  2040
           ||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981  VRDYCKYEKLGDHIFVTAVSSPNVCFTQVPPTLRAAVADGVVQVCYLGEPKTPWTTSAC  2040

Qy   2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE  2100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE  2100

Qy   2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL  2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL  2160

Qy   2161  QRSDPLEGPSNLPPSPPVLQAMPMLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK  2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161  QRSDPLEGPSNLPPSPPVLQAMPMLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK  2220

Qy   2221  EVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221  EVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280

Qy   2281  TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281  TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340

Qy   2341  VRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341  VRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400

Qy   2401  GEIPSHYRQTVIVPKEEVFVKTPQKPTKKKPPRLLISYPHLEMRCVEKMYGQVAPDVVKAV  2460
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401  GEIPSHYRQTVIVPKEEVFVKTPQKPTKKKPPRLLISYPHLEMRCVEKMYGQVAPDVVKAV  2460

Qy   2461  MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461  MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520

Qy   2521  RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK  2580
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521  RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK  2580

Qy   2581  NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN  2640
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581  NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN  2640
```

```
Qy   2641 VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHYPCLWVSRVLAVHF 2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHYPCLWVSRVLAVHF 2700

Qy   2701 MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL 2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL 2760

Qy   2761 TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN 2820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN 2820

Qy   2821 YCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS 2864
           ||||||||||||||||||||||||||||||||||||||||||||
Db   2821 YCDVYSPEGDVFTPQRRLQKFLVKYLAVIVFALGLIAVGLAIS 2864
```

```
Query Match          99.6%;    Score 15317;    DB 2;    Length 2864;
Best Local Similarity 99.6%;    Pred. No. 0;
Matches 2852; Conservative  0; Mismatches  8; Indels  0; Gaps  0;

Qy     1 MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD 60

Qy    61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPV 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPV 120

Qy   121 CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180

Qy   181 CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240

Qy   241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK 300

Qy   301 LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI 360

Qy   361 AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF 420
          ||||||||||||||||||||||||||||||||: |||||||||||||||||||||||||
Db   361 AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPVTLEYNNSISWYPYTIPGARGCMVKF 420

Qy   421 KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG 480

Qy   481 SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ 540

Qy   541 GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY 600

Qy   601 LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFTCCYLRCRLRYAAL 660

Qy   661 LGFVPMAAGLPLTFFVAAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALLVGPWPLVAL 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 LGFVPMAAGLPLTFFVAAAAAQPDYDWWVRLLVAGLVLWAGRDRGPRIALLVGPWPLVAL 720

Qy   721 LTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQRWENWF 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 LTLLHLATPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQRWENWF 780

Qy   781 WNVTLRPERPFLVLVCFPGATYDALVTFCVCHVALLCLTSSAASFGTDSRVRAHRMLVR 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 WNVTLRPERFLLVLVCFPGATYDTLVTFCVCHVALLCLTSSAASFGTDSRVRAHRMLVR 840

Qy   841 LGKCHAWYSHYLVKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQOBPFFPFEGKARVY 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 LGKCHAWYSHYLVKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPFEGKARVY 900

Qy   901 RNEGRRLACGDTVDGLVVARLGDLVPAGLAMPPDGWAITAPFTLQCLSERGTLSAMAVV 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 RNEGRRLACGDTVDGLPVVARLGDLVPAGLAMPPDGWAITAPFTLQCLSERGTLSAMAVV 960

Qy   961 MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020
```

```
RESULT 3
AAR82072
ID    AAR82072 standard; protein; 2864 AA.
XX
AC    AAR82072;
XX
DT    27-AUG-2003  (revised)
DT    02-JUL-1996  (first entry)
XX
DE    Hepatitis GB virus (HGBV) clone GB contig B protein prod.
XX
KW    Hepatitis GB virus; HGBV; diagnosis; treatment; vaccine; reagents; non-A;
KW    non-B; non-C; non-D; non-E; clone; GB contig B; tamarin; infected plasma;
KW    lambda phage; cDNA library.
XX
OS    Hepatitis G virus.
XX
PN    WO9521922-A2.
XX
PD    17-AUG-1995.
XX
PF    14-FEB-1995; 95WO-US002118.
XX
PR    14-FEB-1994; 94US-00196030.
PR    13-MAY-1994; 94US-00242654.
PR    29-JUL-1994; 94US-00283314.
PR    23-NOV-1994; 94US-00344185.
PR    23-NOV-1994; 94US-00344190.
PR    27-JAN-1995; 95US-00344557.
XX
PA    (ABBO ) ABBOTT LAB.
XX
PI    Simons JN, Pilot-Matias TJ, Dawson GJ, Schlauder GG, Desai SM;
PI    Leary TP, Muerhoff AS, Erker JC, Buijk SL, Mushahwar IK;
XX
PT    Non-A, non-B, non-C, non-D, non-E Hepatitis virus reagents - useful for
PT    diagnosis and therapy of hepatitis GB virus.
XX
PS    Example 9; Page 434-447; 661pp; English.
XX
CC    Double stranded hepatitis GB virus (HGBV) DNA obtd. from HGBV infected
CC    tamarin plasma, using standard procedures, was used to prepare a lambda
CC    phage HGBV cDNA library. Clones were rescued from the lambda phage,
CC    searched against a sequence database and found to be unique HGBV
CC    sequences. The clones were then used to assemble the sequences
CC    AAT00129/30 (GB contig A and B) which encode the proteins AAR94345-47
CC    (the 3 possible coding strand reading frames) and AAR82072, respectively.
CC    Reagents which comprise the HGBV DNA, or its protein prods. can be used
CC    for the diagnosis, therapy or in a vaccine to prevent HGBV infection.
CC    (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ    Sequence 2864 AA;
```

WPI; 1995-293123/38.
DR   N-PSDB; AAT00130.

```
Qy  1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG  1080

Qy  1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ  1140

Qy  1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN  1200

Qy  1201  TGASLITYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  TGASLITYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT  1260

Qy  1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA  1320

Qy  1321  RKGITAVSYYRRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  RKGITAVSYYRRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT  1380

Qy  1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYG  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYG  1440

Qy  1441  LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ  1500

Qy  1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT  1560

Qy  1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEECASFIP  1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEECASFIP  1620

Qy  1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIIEYCCGLVTLPDNFFA  1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIIEYCCGLVTLPDNFFA  1680

Qy  1681  SCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  SCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF  1740

Qy  1741  VFDMLGGYAAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  VFDMLGGYAAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT  1800

Qy  1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE  1860

Qy  1861  DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ  1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ  1920

Qy  1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDTDWTSLVVNYG  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDTDWTSLVVNYG  1980

Qy  1981  VRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDGVQVQCYLGEPKTPWTTSAC  2040
          |||||||||::|||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  VRDYCKYEKLGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDGVQVQCYLGEPKTPWTTSAC  2040

Qy  2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE  2100
          |||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db  2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSINNTPSDEAAVSALVFKQE  2100

Qy  2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPRSVPGVSCPESL  2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPRSVPGVSCPESL  2160

Qy  2161  QRSDPLEGPSNLPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK  2220
          ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  QRSDPLEGPSNLPSSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK  2220

Qy  2221  EVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280
          ||||||||:|||||:|||||:||||||||||||||||||||||||||||||||||||||
Db  2221  EVSEWSDGSWSTTTTASSVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280

Qy  2281  TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340

Qy  2341  VRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  VRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400

Qy  2401  GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV  2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV  2460

Qy  2461  MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520

Qy  2521  RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK  2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK  2580

Qy  2581  NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN  2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN  2640

Qy  2641  VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF  2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF  2700

Qy  2701  MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL  2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL  2760

Qy  2761  TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN  2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN  2820

Qy  2821  YCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS  2864
          ||||||||||||:|||||||||||||||||||||||||||||||
Db  2821  YCDVYSPEGDVFVTPQRRLQKFLVKYLAVIVFALGLIAVGLAIS  2864
```

```
XX
PR   14-FEB-1994;    94US-00196030.
PR   13-MAY-1994;    94US-00242654.
PR   29-JUL-1994;    94US-00283314.
PR   23-NOV-1994;    94US-00344185.
PR   23-NOV-1994;    94US-00344190.
PR   30-JAN-1995;    95US-00377557.
XX
PA   (ABBO )  ABBOTT LAB.
XX
PI   Dawson GJ, Leary TP, Muerhoff AS,  Pilot-Matias TJ, Buijk SL;
PI   Mushahwar IK, Simons JN, Desai SM,  Erker JC,  Schlauder GG;
XX
DR   WPI; 2000-338307/29.
XX
PT   Detecting target hepatitis GB virus nucleic acid in a test sample
PT   suspected of containing HGBV comprises reacting the test sample the HGBV
PT   polynucleotide probe and detecting the complex that contains target HGBV.
XX
PS   Claim 1; Col 437-452; 369pp; English.
XX
CC   The present invention describe a method for detecting target hepatitis GB
CC   virus (HGBV) nucleic acid (THN) in a test sample (T) suspected of
CC   containing HGBV. The method involves reacting (T) with a HGBV
CC   polynucleotide probe (I) containing 15 contiguous nucleotides, and which
CC   selectively hybridies to the HGBV genome or its full complement, and
CC   detecting the complex that contains THN, indicating the presence of
CC   target HGBV. The method is used for detecting target HGBV nucleic acid in
CC   the test sample suspected of containing HGBV and for characterisation of
CC   newly ascertained etiological agent of non-A, non-B, non-C, non-D and non
CC   -E hepatitis causing agents collectively termed as hepatitis GB virus.
CC   AAA55270 to AAA55489 and AAB08985 to AAB09480 represent nucleotide and
CC   protein sequences used in the exemplification of the present invention.
CC   (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ   Sequence 2864 AA;
```

```
Query Match             99.6%;  Score 15317;  DB 3;  Length 2864;
Best Local Similarity   99.6%;  Pred. No. 0;
Matches 2852;  Conservative 4;  Mismatches 4;  Indels 8;  Gaps 0;
```

```
Qy      1  MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRARCRKVQRDARPNYKIAGIHD  60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRARCRKVQRDARPNYKIAGIHD  60

Qy     61  GLQTLAQAALPAHGWGWRQDPRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPV 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  GLQTLAQAALPAHGWGWRQDPRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPV 120

Qy    121  CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180

Qy    181  CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG  240

Qy    241  ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK  300

Qy    301  LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI  360

Qy    361  AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF  420
           |||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db    361  AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPVTLEYNNSISWYPYTIPGARGCMVKF  420

Qy    421  KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG  480
```

```
Qy    481  SKEMFKPHNWMSGHLYPEGSDTPIVVFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  SKEMFKPHNWMSGHLYPEGSDTPIVVFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ  540

Qy    541  GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY  600

Qy    601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL  660

Qy    661  LGFVPMAAGLPLTFFVAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALLVGPWPLVAL  720
           |||||||||||||||||||||||:|||||||||||||||||||||||||:|||||||||
Db    661  LGFVPMAAGLPLTFFVAAAAQPDYDWWVRLLVAGLVLWAGRDRGPRIALLVGPWPLVAL  720

Qy    721  LTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLMQRWENWF  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  LTLLHLATPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLMQRWENWF  780

Qy    781  WNVTLRPERFFLVLVCFPGATYDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  WNVTLRPERFLLVLVCFPGATYDTLVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR  840

Qy    841  LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPEGKARVY  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPEGKARVY  900

Qy    901  RNEGRRLACGDTVDGLPVVARLGDLVPAGLAMPPDGWAITAPPTLQCLSERGTLSAMAVV  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  RNEGRRLACGDTVDGLPVVARLGDLVPAGLAMPPDGWAITAPPTLQCLSERGTLSAMAVV  960

Qy    961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020

Qy   1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG 1080

Qy   1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQTAHATLDTKPTVPNEYSVQ  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQTAHATLDTKPTVPNEYSVQ  1140

Qy   1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200

Qy   1201  TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATVLGIGKVLTEAPSKNVRLVVLAT  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATVLGIGKVLTEAPSKNVRLVVLAT  1260

Qy   1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA 1320

Qy   1321  RKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  RKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT 1380

Qy   1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYVVDGSCTPSGMVPECNIVEAPDAAKAWYG 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYVVDGSCTPSGMVPECNIVEAPDAAKAWYG 1440

Qy   1441  LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441  LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ 1500

Qy   1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT 1560

Qy   1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVDEEIVEECASFIP  1620
```

```
Db
Qy  1561 SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEECASFIP 1620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  1621 LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA 1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  1681 SCVFAFTAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF 1740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  1741 VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVGVLSACAMFALTT 1800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  1801 AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE 1860
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  1861 DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ 1920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  1921 ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVNYG 1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  1981 VRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDGVQVQCYLGEPKTPWTTSAC 2040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  2041 CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE 2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  2101 LRRTNQLLEAISAGVDTTKLPAPSIEEVVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL 2160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  2161 QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK 2220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  2221 EVSEWSDESWSTATTASSYVTGPPYPYKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW 2280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  2281 TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  2341 VRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA 2400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  2401 GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV 2460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  2461 MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH 2520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  2521 RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK 2580
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  2581 NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCCSN 2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db
Qy  2641 VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF 2700
```

```
Db  2641 VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF 2700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2701 MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL 2760
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL 2760

Qy  2761 TDMTMPPLRAWRKKARAVLASAKRRGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN 2820
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN 2820

Qy  2821 YCDVYSPEGDVFTTPQRRLQKFLVKYLAVIVFALGLIAVGLAIS 2864
         ||||||||||| |||||||||||||||||||||||||||||||
Db  2821 YCDVYSPEGDVFVTPQRRLQKFLVKYLAVIVFALGLIAVGLAIS 2864
```

RESULT 5

```
AAB09038
ID   AAB09038 standard; protein; 1422 AA.
XX
AC   AAB09038;
XX
DT   06-AUG-2003 (revised)
DT   30-AUG-2000 (first entry)
XX
DE   Hepatitis GB virus protein sequence SEQ ID NO:83.
XX
KW   Hepatitis GB virus; HGBV; diagnosis; therapeutic; immunogenic; infection;
KW   detection; characterisation; hepatitis.
XX
OS   Hepatitis GB virus.
XX
PN   US6051374-A.
XX
PD   18-APR-2000.
XX
PF   07-JUN-1995;   95US-0048845.
XX
PR   14-FEB-1994;   94US-00196030.
PR   13-MAY-1994;   94US-00242654.
PR   29-JUL-1994;   94US-00283314.
PR   23-NOV-1994;   94US-00344185.
PR   23-NOV-1994;   94US-00344190.
PR   30-JAN-1995;   95US-00377557.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Dawson GJ, Leary TP, Muerhoff AS, Pilot-Matias TJ, Buijk SL;
PI   Mushahwar IK, Simons JN, Desai SM, Erker JC, Schlauder GG;
XX
DR   WPI; 2000-338307/29.
XX
PT   Detecting target hepatitis GB virus nucleic acid in a test sample
PT   suspected of containing HGBV comprises reacting the test sample the HGBV
PT   polynucleotide probe and detecting the complex that contains target HGBV.
XX
PS   Example 5; Col 223-230; 369pp; English.
XX
CC   The present invention describe a method for detecting target hepatitis GB
CC   virus (HGBV) nucleic acid (THN) in a test sample (T) suspected of
CC   containing HGBV. The method involves reacting (T) with a HGBV
CC   polynucleotide probe (I) containing 15 contiguous nucleotides, and which
CC   selectively hybridises to the HGBV genome or its full complement, and
CC   detecting the complex that contains THN, indicating the presence of
CC   target HGBV. The method is used for detecting target HGBV nucleic acid in
CC   the test sample suspected of containing HGBV and for characterisation of
CC   newly ascertained etiological agent of non-A, non-B, non-C, non-D and non
CC   -E hepatitis causing agents collectively termed as hepatitis GB virus.
CC   AAA55270 to AAA55489 and AAB08985 to AAB09480 represent nucleotide and
CC   protein sequences used in the exemplification of the present invention.
CC   (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ   Sequence 1422 AA;
```

```
Query Match       48.0%;  Score 7380;  DB 3;  Length 1422;
Best Local Similarity 97.3%;  Pred. No. 0;
Matches 1392;  Conservative 2;  Mismatches 27;  Indels 10;  Gaps 2;

Qy 1005 AHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYW 1064
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 AHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYW 60

Qy 1065 CVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAH 1124
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVGQIRVRPLVCAGYHPQYTAH 120

Qy 1125 ATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHA 1184
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQXKXEVLVLNPSVATTASMPKYMHA 180

Qy 1185 TYGVNPNCYFNGKCINTGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKV 1244
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 TYGVNPNCYFNGKCINTGASLTYSTYGMYLTGRCSRNYDVIICDECHATDRTTVLGIGKV 240

Qy 1245 LTEAPSKNVRLLVVLATATPGVIPTPHANITEIQLTDEGTIPFHGKKIKENLKKGRHLI 1304
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 LTEAPSKNVRLLVVLATATPGVIPTPHANITEIQLTDEGTIPFHGKKIKENLKKGRHLI 300

Qy 1305 FEATKCHCDELANELARKGITAVSYYRGCDISKIPEGDCVVATDALCTGYTGDFDSVYD 1364
        |||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db  301 FEATKCHCDELANELARKGITAVSYYRGCDISKMPBGDCVVATDALCTGYTGDFDSVYD 360

Qy 1365 CSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVP 1424
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 CSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVP 420

Qy 1425 ECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVN 1484
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 ECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVN 480

Qy 1485 TAKRTADNYVLLTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPE 1544
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TAKRTADNYVLLTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGCDACPGPE 540

Qy 1545 PSEVTRYQMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAP 1604
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 PSEVTRYQMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAP 600

Qy 1605 VVDEEEIVEECASFIPLEAMVAAIDKLKSTITTTSPTTLETALEKLNTFLGPHAATTLAI 1664
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 VVDEEEIVEECASFIPLEAMVAAIDKLKSTITTTSPTTLETALEKLNTFLGPHAATTLAI 660

Qy 1665 IEYCCGLVTLPDNPASCVRFAFIAGITTPLPHKIKMPLSLFGGAIASKLTDARGALAFMM 1724
        |||||||||||||||||||||||||||||||||||||||||||||:||||||:||||||
Db  661 IEYCCGLVTLPDNPASCVRFAFIAGITTPLPHKIKMPLSLFGGAIASKLTDARXALAFMM 720

Qy 1725 AGAAGTALGTWTSVGFVFDMLGGYAAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAA 1784
        |||||||||||||||||||:||||||:||||||||||||||||:|||||||||||||||
Db  721 AGAXGTALGTWTSVGFVFDMLGGYAGASSTACLTFKCLMGEWXTMDQLAGLVYSAFNPAA 780

Qy 1785 GVGVLSACAMFALTTAGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTP 1844
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GVGVLSACAMFALTTAGPDHWPNRLLTMLARSNTVCXEYFIATRDIRRKILGILEASTP 840

Qy 1845 WSVISACIRWLHTPEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYS 1904
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 WSXISACIRWLHTPEDDCGLIAWGLXIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYS 900

Qy 1905 CQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGS 1964
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 CQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGS 960

Qy 1965 ARPDPTDWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVADGVQV 2024
        |||||||||||||||||||||||:|||:||||||||||||||||||||||||||||||
Db  961 ARPDPTDWTSLVVNYGVRDYCKYEKLGDHIFVTAVSSPNVCFTQVPPTLRAAVADRVQV 1020

Qy 2025 QCYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNN 2084
        |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db 1021 QXYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGGRMQLNLRDRLEANDCNSINN 1080

Qy 2085 TPSDEAAVSALVFKQELRRTNQLLEAISAGVDTTKLPAPS-IEEVVVRKRQPRARTGSLT 2143
        |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db 1081 TPSDEAAVSALVFKQELRRTNQLLEAISAGVDTTKLPAPSQIEEVVVRKRQPRARTGSLT 1140

Qy 2144 LPPPPRSVPGVSCPESLQRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAM 2203
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 LPPPPRSVPGVSCPESLQRSDPLEGPSXLPSSPPVLQLAMPMPLLGAGECNPFTAIGCAM 1200

Qy 2204 TETGGGPDDLPSYPPKKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPT 2263
        |||||||||||||||||||||||||||||||||||||||||||||||||||||::||:||
Db 1201 TETXGPXPXLPSYPPKKEVSEWSDESWSTTTASSYVTGPPYPKIRGKDSTQSATAKRPT 1260

Qy 2264 KKKLGKSEFSCSMSYTWTDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQ 2323
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 KKKLGKSEFSCSMSYTWTDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQ 1320

Qy 2324 KVTINRQPLFPPSYHKQVRLAKEKASKVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVR 2383
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 KVTINRQPLFPPSYHKQVRLAKEKASKVGVMWDYDEVAAHTPSKSAKSHITGLRGTD-- 1378

Qy 2384 SGAARKAVLDLQKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLI 2434
        ------|||||||||||||||||||||||||||||||||||||||||||
Db 1379 ------VLDLQKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLI 1422
```

RESULT 6
AAR82068
ID AAR82068 standard; protein; 1422 AA.
XX
AC AAR82068;
XX
DT 27-AUG-2003 (revised)
DT 02-JUL-1996 (first entry)
XX
DE Hepatitis GB virus (HGBV) clone protein prod.
XX
KW Hepatitis GB virus; HGBV; diagnosis; treatment; vaccine; reagents; non-A;
KW non-B; non-C; non-D; non-E; clone; tamarin; infected plasma;
KW lambda phage; cDNA library.
XX
OS Hepatitis G virus.
XX
FH Key   Location/Qualifiers
FT Misc-difference 1..1422
FT   /note= "others correspond to degenerate or STOP codons in
FT   AAT00052"
XX
PN WO9521922-A2.
XX
PD 17-AUG-1995.
XX
PF 14-FEB-1995; 95WO-US002118.
XX
PR 14-FEB-1994; 94US-00196030.
PR 13-MAY-1994; 94US-00242654.
PR 29-JUL-1994; 94US-00283314.
PR 23-NOV-1994; 94US-00344185.
PR 23-NOV-1994; 94US-00344190.
PR 27-JAN-1995; 95US-00344557.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Simons JN, Pilot-Matias TJ, Dawson GJ, Schlauder GG, Desai SM;
PI Leary TP, Muerhoff AS, Erker JC, Buijk SL, Mubahwar IK;
XX
DR WPI; 1995-293123/38.
DR N-PSDB; AAT00052.
XX
PT Non-A, non-B, non-C, non-D, non-E Hepatitis virus reagents - useful for
PT diagnosis and therapy of hepatitis GB virus.

```
XX
PS   Example 5; Page 245-249; 661pp; English.
XX
CC   Double stranded hepatitis GB virus (HGBV) DNA obtd. from HGBV infected
CC   tamarin plasma, using standard procedures, was used to prepare a lambda
CC   phage HGBV cDNA library. The cDNA clone AAT00052, which encodes the
CC   proteins AAR82066-71 (the 6 possible reading frames), was rescued from
CC   the lambda phage, searched against a sequence database and found to be an
CC   unique HGBV sequence. Reagents which comprise the HGBV DNA, or its
CC   protein prods. can be used for the diagnosis, therapy or in a vaccine to
CC   prevent HGBV infection. (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 1422 AA;

Query Match             48.0%;   Score 7375;  DB 2;  Length 1422;
Best Local Similarity   97.2%;   Pred. No. 0;
Matches 1391;  Conservative  2;  Mismatches  28;  Indels  10;  Gaps  2;

Qy  1005  AHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYW 1064
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  AHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYW 60

Qy  1065  CVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAH 1124
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db    61  CVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVGQIRVRPLVCAGYHPQYTAH 120

Qy  1125  ATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHA 1184
          |||||||||||||||||||||||||||||||| ||||| |||||||||||||||||||||
Db   121  ATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQXKEVLVLNPSVATTASMPKYMHA 180

Qy  1185  TYGVNPNCYFNGKCTNTGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKV 1244
          |||||||||||||||||||||||||||||||| |||||||||||||||||| ||||||||
Db   181  TYGVNPNCYFNGKCTNTGASLTYSTYGMYLTGRCSRNYDVIICDECHATDRTTVLGIGKV 240

Qy  1245  LTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDEGTIPHGKKIKEENLKKGRHLI 1304
          |||||||||||||||||||||||||||||||||| :||||||||||||||||||||||||
Db   241  LTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDEGTIPHGKKIKEENLKKGRHLI 300

Qy  1305  FEATKKHCDELANELARKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYD 1364
          ||||||||||||||||||||||||||||||| :|||||||||||||||||||||||||||
Db   301  FEATKKHCDELANELARKGITAVSYYRGCDISKMPEGDCVVVATDALCTGYTGDFDSVYD 360

Qy  1365  CSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVP 1424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  CSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVP 420

Qy  1425  ECNIVEAFDAAKAWYGLSSTEAQTILDYRTQPGLPAIGANLDEWADLFSMVNPEPSFVN 1484
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  ECNIVEAFDAAKAWYGLSSTEAQTILDYRTQPGLPAIGANLDEWADLFSMVNPEPSFVN 480

Qy  1485  TAKRTADNYVLLTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPE 1544
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   481  TAKRTADNYVLLTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGCDACPGPE 540

Qy  1545  PSEVTRYQMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAP 1604
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  PSEVTRYQMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAP 600

Qy  1605  VVDEEIVEECASFIPLEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATTILAI 1664
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  VVDEEIVEECASFIPLEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATTILAI 660

Qy  1665  IEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMM 1724
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   661  IEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARXALAFMM 720

Qy  1725  AGAGTALGTWTSVGPVVFDMLGGYAAASTACLTFKCLMGEWPTMDQLAGLVVSAFNPAA 1784
          |||| |||||||||| |||||||||| ||||||||||||||| |||||||||||||||||
Db   721  AGAXGTALGTWTSVGFVVFDMLGGYAGASSTACLTFKCLMGEWXTMDQLAGLVVSAFNPAA 780

Qy  1785  GVVGVLSACAMFALTTAGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTP 1844
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   781  GVVGVLSACAMFALTTAGPDHWPNRLLTMLARSNTVCKEYFIATRDIRRKILGILEASTP 840

Qy  1845  WSVISACIRWLHTPTEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYS 1904
          ||| |||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   841  WSXISACIRWXHTPTEDDCGLIAWGLXIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYS 900

Qy  1905  CQKGKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGS 1964
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  CQKGKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGS 960

Qy  1965  ARPDPTDWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDGVQV 2024
          |||||||||||||||||||||| :|||||||||||||||||||||||||||||||||:|||
Db   961  ARPDPTDWTSLVVNYGVRDYCKYEKLGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDRVQV 1020

Qy  2025  QCYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNN 2084
          | :|||||||||||||||||||||||||||||||||||:|||||||||| | | |||:||
Db  1021  QXYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGGRMQLNLRDRLEANDCNSINN 1080

Qy  2085  TPSDEAAVSALVFKQELRRTNQLLEAISAGVDTTKLPAPS-IEEVVRKRQFRARTGSLT 2143
          |||||||||||||||||||||||||||||||||||||||| :|||||||||||||||||
Db  1081  TPSDEAAVSALVFKQELRRTNQLLEAISAGVDTTKLPAPSQIEVVVRKRQFRARTGSLT 1140

Qy  2144  LPPPRRSVPGVSCPESLQRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAM 2203
          |||||||||||||||||||||||||| |||:|||||||||||||||||||||||||||||
Db  1141  LPPPRRSVPGVSCPESLQRSDPLEGPSXLPSSPPVLQLAMPMPLLGAGECNPFTAIGCAM 1200

Qy  2204  TETGGGPDDLPSIYPPKKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPT 2263
          ||| :| ||| |||||||||||||||||| ||||||||||||||||||||||||| |||:|
Db  1201  TETKGXPXXLPSYPPKKEVSEWSDESWSTTTTASSYVTGPPYPKIRGKDSTQSATAKRPT 1260

Qy  2264  KKKLGKSEFSCSMSYTWTDVISFKTASKVLSATRAITSGFLKQRSLVVTEPRDAELRKQ 2323
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  KKKLGKSEFSCSMSYTWTDVISFKTASKVLSATRAITSGFLKQRSLVVTEPRDAELRKQ 1320

Qy  2324  KVTINRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVR 2383
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  KVTINRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTD-- 1378

Qy  2384  SGAARKAVLDLQKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLI 2434
                ||||||||||||||||||||||||
Db  1379  ------VLDLQKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLI 1422


RESULT 7
AAB59172
ID   AAB59172 standard; protein; 3033 AA.
XX
AC   AAB59172;
XX
DT   21-MAR-2001  (first entry)
XX
DE   Protein encoded by infectious Hepatitis C virus 2a genotype.
XX
KW   GBV-B; hepatitis C virus; HCV; vaccine.
XX
OS   Hepatitis C virus.
XX
PN   WO200075337-A1.
XX
PD   14-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015293.
XX
PR   04-JUN-1999;  99US-0137694P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Bukh J, Yanagi M, Emerson SU, Purcell RH;
XX
DR   WPI; 2001-091214/10.
XX
PT   New infectious nucleic acids of the GB virus-B clone, useful for
PT   indirectly studying the molecular properties of hepatitis C virus (HCV)
PT   and in developing vaccines and therapeutics for HCV.
```

```
XX
PS    Disclosure; Page 82-94; 96pp; English.
XX
CC    The present invention relates to GB virus-B. The nucleic acid molecules
CC    of the invention are useful for indirectly studying the molecular
CC    properties of hepatitis C virus (HCV). The infectious nucleic acid
CC    sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used
XX    in the development of vaccines and therapeutics for HCV
SQ    Sequence 3033 AA;

Query Match      21.3%;      Score 3277.5;    DB 4;    Length 3033;
Best Local Similarity    30.3%;    Pred. No. 1.8e-244;
Matches  987;    Conservative 436;    Mismatches 1197;    Indels 641;    Gaps 117;
```

```
Qy     8  TSPVPAPRTRKNKQTQASYPVSIK----------TSVERGQORARKRVQRDARPR          51
          :: :  :::     ::     :    :::|                  ||::|
Db     3  TNPKPQRKTKNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR     59

Qy    52  NYKIAGIHDGLQTLAQAALPAH----------------------GWGRQDPRHKSRNL      87
          : :||     :   :  ::                         ||  ||:|  ||||
Db    60  GRRQPIPKDRRSTGKSWGKPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPNDPRHRSRNV    119

Qy    88  GILLDYPLGWIGDVTTHTPLVG-PLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLF    143
          | :: | : |::||:   :|   | || |     ::|  :|||||||:||||
Db   120  GKVIDTLTCGFADLMGYIPVVGAPL--GGVARALAHGVRVLEDGVNFATGNLPGCSFSIF    177

Qy   144  VVCLLS-LACPCSGARVTDPDTNTTLTLNCCQRNQVIYCSPSTCLHEPGCVIC------AD  197
          :::| | | | ||:||:|    |    :| |:|:::|   ::  |||||:|::|       :
Db   178  LLALLSCITTPVSAAEVKNISTG-YMVTNDCTNDSITWQLQAAVLHVPGCVPCEKVGNAS    236

Qy   198  ECWVPANPYIS--HPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV-   254
          :||:| :|  :  :| :   :    :||||:|:|: |||:  |::| |||  :|   :|
Db   237  QCWIPVSPNVAVQRPGALT---QGLRTHIDMVVMSATLCSALYVGDLCGGVMLAAQMFIV    293

Qy   255  ---RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTKLASQVPYAIAT   311
              ::| : :|::|  |  ||   | |  :  :|  |:|  |  :  : |  :||| : |
Db   294  SPQHHWFVQ------DCNCSIYPGTITGHRMAWDMMMNWSPTATMILAYAMRVPEVIID     346

Qy   312  MFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPT-GCSIAE-------   362
           ::  |: | || | :::|   |    |:::    ::|  |:|    ||  |
Db   347  IISGAHWGVMFGLAYPSMQGAWAKVVVILLL----AAGVDARTHTVGGSAAQTTGRLTSL    402

Qy   363  ------------------------FCSPLMIP-------CP-----CHS------      375
                                  | || :|       ||     ||:
Db   403  FDMGPRQKIQLVNTNGSWHINRTALNCNDSLHTGFIASLFYTHSFNSSGCPERMSACRSI    462

Qy   376  ----------YLSENVSE-----------------VICYSPKWTRPITLE          398
                    :| ||:                    ||  :|||  | | |
Db   463  EAFRVGWGALQVEDNVTNPEDMRPYCWHYPPRQCGVVSAKTVCGPVYCFTPS---PVVVG    519

Qy   399  YNNSISWYPYTIPGARGCMVFKNNT--------WGC------------CRIR------N   432
             | : | ||  | |||  : |::         |||             ||:       |
Db   520  TTDRLGAPTYTW-GENETDVFLLNSTRPPLGSWFGCTWMNSSGYKTICGAPPCRTRADFN    578

Qy   433  VPSYCTMGTDAVWNDTRNTYEACGVTPWLT---------TAWHNGSALKLAILQ---YPG   480
          |:  |  ||||:|:||||||  ||:|||:          ||||||||||| ||:   ||
Db   579  ASTDLLCPTDCFRKHPDTTYLKCGSGPWLTPRCLIDYPYRLWHYPCTVNYTIFKIRMYVG    638

Qy   481  SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLP---PERWARLPGTPPVVRGSWL   536
           | :::| | :|||||:|:|||||||||||||||| |    ||:     ||||||||
Db   639  GVE-----HRLTAACNFTRGDRCNL--EDRDRSQLSPLLHSTTEWAILPCS--------    682

Qy   537  QVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYL   596
           :    |||| :| :||: | ::  |  |:  :  ||  |||:::||||:|||:  :||
Db   683  ------YSDLPALSTGLLHLHQNIVDVQFMYGLSPALTKYIVRWEWVILLFLLLADAR--    734

Qy   597  ILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLR   656
           :| | ||:|     ||   ||| || |::::| ||:|||||||||||||||||||||||
Db   735  VCACLWMLIL------------------------LGQAEAALEKIVILHAA------S     762

Qy   657  YAALLGFVPMAAGLPLTFFVAAAAQPDYDWMVRLLVAGLVLWAGRNRGHRIALLVGPWP   716
           :|| |||||:||||||||||||||||||||||||||||||||||||||||:||||||
Db   763  AASCNGFLYF-----------VIFFVAA-------------WYIKGRVVPLATYS-    799
```

```
Db   763  AASCNGFLYF-----------VIFFVAA-------------WYIKGRVVPLATYS-    LTGLWS   799
Qy   717  LVALLTLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQRW            776

Db   800  F--SLLLIALPQQAYAYDASVHGQIGAALLVMTLFLTPGYKTLLSR------FLW-W      849
Qy   777  -------ENWFMNVTLRPER---FFLJLVCFPGATYDALVTFCVCHVALLCLTSS       821

Db   850  LCYLLTLGEAMVQEWAPPMQVRGGRDGIIWAVAIFYPGVVFDITKW------LLAVLGP    902
Qy   822  AASFFGTDSR----VRAHRMLVRLGKC---HAWYSHYVLKFFLLVFGE-NGVFFYKHL-    871

Db   903  AYLLKGALTRVPYFVRAHAL--LRMCTMARHLAGGKRYV-QMALLALGRWTGTYIYDHLT    958
Qy   872  -HGDVLPNDFASKLPLQEP--FPPFEGKARVYRNEGRRLACGDTVDGLPVARLGDLVFA    928

Db   959  PMSDWAASGLRDLAVAVEPIIFSPMEKKVIVWGAE--TAACGDILHGLPVSARLGREVLL   1016
Qy   929  GLAMPPD------GWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSY  983

Db  1017  G----PADGYTSKGWSLLAPITAYAQQTRGLLGTIVVSMTGRDKTEQAGEIQVLSTVTQSF 1073
Qy   984  MGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCSCGETK  1043

Db  1074  LGTSISGVLWTVYHGAGNKTLAGSRGPVTQMYSSAEGDLVGWPSPPGTKSLEPCTCGAVD  1133
Qy  1044  GYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSG-  1102

Db  1134  LYLVTRNADVIPARRRGDKRGALLSPRPLSTLKGSSGGPVLCPRGHAVGVFRAAVCSRGV  1193
Qy  1103  --SVSQIRVRPLVCAGYHPQTAHATLDTKPTVPNEYSVQILLAPTGSGKSTKLPLSYMQ  1160

Db  1194  AKSIDFIPVETLDIVTRSPTFSDNS---TPPAVPQTYVGYLHAPTGSGKSTKVPVAYAA  1250
Qy  1161  EKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYLT-GACS  1219

Db  1251  QGYKVLVLNPSVAATLGFGAYLSKAHGINPNIRTGVRTVTTGAPITYSTYGKFLADGGCA  1310
Qy  1220  RN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANITEIQ  1278

Db  1311  GGAYDIIICDECHAVDSTTILGIGTVLDQAETAGVRLTVLATATPPGSVTTPHPNIEEVA  1370
Qy  1279  LTDEGTIPFHGKKIKEENLKKGRHLIFEATKCHCDELANELARKGITAVSYYRGCDISKI  1338

Db  1371  LGQEGEIPFYGRAIPLSYIKKGGRHLIFCHSKKKCDELAAALRGMGLNSVAYYRGLDVSVI 1430
Qy  1339  P-EGDCVVVATDALCTGYTGDPDSVYQCLSMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQ  1397

Db  1431  PTQGDVVVVATDALMTGYTGDFDSVIDCNVAVTQVVDFSLDPTFTITQIVPQDAVSRSQ   1490
Qy  1398  RRGRTGRGRAGIYYYVDGSCTPSGSMVPECNIVEAFDAKAWYGLSSTEAQTILDTYRTQP  1457

Db  1491  RRGRTGRGRLGIYRYVSTGERASGMFDSVVLCECYDAGAAWYELFPSETTVRLRAYFNTP  1550
Qy  1458  GLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNVVLLTAAQLQLCHQGYAAPNDAP  1516

Db  1551  GLPVCQDHLEFWEAVFTGLTHIDAHFLSQTKQSGENFAYLTAYQATVCAR---AKAPPP  1606
Qy  1517  RWQGARLGKKPCCGVLWRLDGADAC-------------PGPEP-------SEVTRY--   1551

Db  1607  SWD-------------VMWK------CLTRLKPTLVGPTPLLYRLGSVTNEVTLTHPVTKYIA 1650
Qy  1552  --QMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDBE  1609

Db  1651  TCMQADLEVMTS-TWVLAGGVLAAVAAYCLAT-GCVCIIGRLHINQ--RAVVAP--DKE  1703
Qy  1610  EI------VEECASFIPL-EAMVAAIDKLKSTI--------------TTTSPFTLETALEKLN 1651

Db  1704  VLYEAFDEMEECASRAALIEEGQRIAEMLKSKIQGLLQQASKQAQDIQP-TVQASWPKVE  1762
Qy  1652  TFLGPHAATILAIIEYCCGLVTLPDNPFPASCVPAFIAGITTPLPHKIKMFLSLFGGAIAS 1711

Db  1763  QFWAKHMWNFISGIQYLAGLSTLPGNPAVASMMAFSAALTSPLSTSTTILLNILGGWLAS  1822
```

```
Qy   1712  KLTDARGALAFMMAGAGAGTALGTWTSVGF---VFDMLGGYAAAASSTACLTFKCLMGEWPT  1768
                :  :  ::: | ||:: |  :      :     |:   :||  :    :  :  |:|  |
Db   1823  QIAPPAGATGFVVSGLVCAAVG---SIGLGKVLVDILAGYGAGISGALVAFKIMSGEKPS  1879

Qy   1769  MDQLAGLVYSAFNPAAGVVGVLSACAMFALTTAGPD----HWPNRLLTMLARSNTVCNEY  1824
              |:: :|: ||||  |  ||||||  |       :|| ||     :| |  | : | :|
Db   1880  MEDVVNLLPGILSPGALVVGVI--CAAILRRHVGPGEGAVQWMNRLLIAFASRGNHVAPTH  1937

Qy   1825  FIATRDIRRKILGILEASTPWSVISACIRWLHTPTEDDCGLI--AWGLEIWQYVCNFFV  1881
             ::  :  |:|| ::       ::   |:|||:    |  :|                   ||
Db   1938  YVTESDASQRVTQLLGSLT----ITSLLRRLHNWITEDCPIPCGGSWLRDVWDWVCTILT  1993

Qy   1882  ICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLY  1941
           |    : :|   | :| ::| |  :  :  | | |:  | ::  |  |   :|     |
Db   1994  DFKNWLTS--KLFPKMPGLPFVSCQKGYKGVWAGTGIMTTRCPCGANISGNVRLGSMRI-  2050

Qy   1942  KGPRTCSNYWRGAVPVNARLCGSARPDT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVS  2000
           :||:|  :  | :|   | |||   |     |:|| | | |  :  :: :|  | : |||
Db   2051  TGPKTCMNIWQGTFPINCYTEGQCVPKPAPNFKVAIWRVAASEYAEVTQHGSYHYITGLT  2110

Qy   2001  SPNVCFTQVP---PTLRAAVADVGVQVQCYLGEPKTPWTTS-ACCYGPDGKGKTVKLPFR  2056
             | : :: :||    :  |    |  :|      | |||    |  |     | :||
Db   2111  TDNL---KVPCQLPSPEFFSWVDGVQIHRFAPTPKPFFRDEVSFCVGLNSFVVGSQLPCD  2167

Qy   2057  VDGHTPGVRMQLNLRD---ALETNDCNSTNNTPSDERAAVSA-LVFKQELRRT-------  2104
              |  ||  : ::|:|      |   |    |||:     |    ::        |
Db   2168  PEPDT-DVLMSM-LTDPSHITAETAARRLARGSPPSEASSSASQLSRATCTTHGKA  2225

Qy   2105  --NQLLEA-ISAGVDTKLPA----------------------PSI-EVVVRKRQFRART  2139
             :|| :|  | :||:| ||                        |  | |::|:| ||:
Db   2226  YDVDMVDANLFMGGDVTRIESGGSKVVVLDSLDPMVEERSDLEPSIPSEYMLPKKRF---  2281

Qy   2140  GSLTLPP--PPRSVPGVSCP--ESLQRSD---------PLEGPSNLPPSPP--------  2177
            | | ||   || | |   |   | | :|            | |  : ||
Db   2282  -----PPALPAWARPDYNPPLVESWKRPDYQPATVAGCALPPPRKTPTPPPRRRRTVGLS  2336

Qy   2178  ----------VLQLAM---PMPLLGAGECNPFTAIGC--------AMTETGG-------  2208
                       |  | |    |||| |  |   | |          :  |
Db   2337  EDSIGDALQQLAIKSFGQPPP---SGDSGLSTGAGAADSGSQTPPDELALSETGSISSMP  2393

Qy   2209  ------GPDDL-------PSYPPKEV-SEWSDE-SWSTATTASSYVTGPPYPKIRGKDST  2254
                 |  :|        || |:|   :|:|   : ::|  :|  :||  | |:
Db   2394  PLEGELGDPDLEPEQVEPQPPPQGGVAAPGSDSGSWSTCSEEDDSVV-----------  2440

Qy   2255  QSAPAKRPTKKKLGKSEFSCSMSYTWTD-VISFKTASKVLSATRAITSGFLKQRSLVVVT  2313
                                  |:| :|    | :  |   |::|  ||: |:|:  | ||
Db   2441  ----------------CCSMSYSWTGALITPCSPEEEKLPINPLSNSLLRYHNKYYCT  2482

Qy   2314  EPRDAELRKQKVTINRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSH  2373
             |  |  :: | :|  :|    |:|||:|   |:|:   : |:|: ||  ||||:|||
Db   2483  TTKSASLRAKKVTFDRMQVLDSYYDSVLKDIKLAASKVTARLLJTMEEACQLTPPHSARSK  2542

Qy   2374  ITGLRGTDVR--SGAARQAVLDLQK-CVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKP  2430
             ||| |:: |   ||  |  |:| :   ::|  | |  |  | :| |  ||| |  |
Db   2543  Y-GFGAKEVRSLSGRAVNHIKSVWKDLLEDSETPI---PTTIMAKNEVFCVDPTKGGKKA  2598

Qy   2431  PRLISYPHLEMRCVEKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMWS--PDAVG  2487
            ||| |  ||  ||  | ||   |  || |  |  |   |  || || | :    || |
Db   2599  ARLIVYPDLGVRVCEKMALYDITQKLPQAVMGASYGFQYSPAQRVEFLLKAWAEKKDPMG  2658

Qy   2488  ATCDTVCFDSTTTPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGY  2547
             |  :||  |  :   |  | ||  : ||  | :||  | ||  | | |  | |||: ||
Db   2659  FSYDTRCFPDSTVTERDIRTEESIYRACSLPEEAHTAIHSLTERLYVGGPMFNSKGQTCGY  2718

Qy   2548  RRCRSSGVYTTSSSNSLTCWLKVNAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAM  2607
           |||  |||  || :  | |:||   || | |   |  |:||| |::|||  | |   |
Db   2719  RRCRASGVLTTSMGNTITCYVKALAACKAAGIIAPTMLVCGDDLVVISESQGTEEDERNL  2778

Qy   2608  RVFASWMKVMGAPQDCVPQPKYISEELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCS  2667
           |  |::|  || |   |    | | ||:||| ::  |     |  ||||||:|  ||  :
Db   2779  RAFTEAMTRYSAPPGDPPRPEYDLELITSCSSNVSVALGPQGRRRYYLTRDPTTPIARAA  2838
```

```
Qy   2668  AEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPV  2727
              ||  | ||: |  |:|::|||  |:|| ||| || ::|  ||  |: ||:|| | |
Db   2839  WETVRHSPVNSWLGNIIQYAPTIWARMVLMTHFFSILMAQDTLDQNLFEMYGAVYSVSP  2898

Qy   2728  EDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRG  2787
           :||  : | |   | ||: | :     |:|||:| |:| || ||: ||:||  :  : ||
Db   2899  LDLPAIIERLHGLDAFSLHTYTPHELTRVASALRKLGAPPLRAWKSRARAVRASLISRGG  2958

Qy   2788  AHAKLARFLL-WHATSR-----PLPDLDKTSVARYTTFNYCDVVSPEGDVFITPQRRLQKF  2842
           |  |||  |  || |:|        ||:|| :     ||:   :   ||:|
Db   2959  RAAVCGRYLFNWAVKTKLKLTPLPEARLLDLSSWFT----VGAGGDIYHSVSRARPRL  3013

Qy   2843  LVKYLAVIVFALGLIAVGLAI  2863
           |   | |:|| |||  |:|:|
Db   3014  L------LFGLLLFVGVGL  3027
```

RESULT 8

```
AAB31168
ID   AAB31168 standard; protein; 3033 AA.
XX
AC   AAB31168;
XX
DT   02-APR-2001 (first entry)
XX
DE   Amino acid sequence of a hepatitis C virus (HCV) clone genotype 2a.
XX
KW   Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;
KW   HCV; vaccine; viral inhibitor; antiviral.
XX
OS   Hepatitis C virus.
XX
PN   WO200075352-A2.
XX
PD   14-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015527.
XX
PR   04-JUN-1999;   99US-0137817P.
XX
PA   (USSH )  US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Nam J,  Bukh J,   Emerson SU,   Purcell RH;
XX
DR   WPI; 2001-071081/08.
DR   N-PSDB; AAC86937.
XX
PT   New nucleic acid comprising a chimeric bovine viral diarrhea virus genome
PT   in which the (non-)structural region has been replaced by hepatitis C
PT   virus (HCV) genome useful for treating or preventing HCV signs and
PT   symptoms.
XX
PS   Disclosure; Page 85-97; 97pp; English.
XX
CC   The specification describes a nucleic acid comprising a chimeric virus
CC   genome, specifically bovine viral diarrhoea virus (BVDV) genome in which
CC   the (non-)structural region has been replaced by the (non-)structural
CC   region of a hepatitis C virus (HCV) genome. The nucleic acids comprising
CC   the chimeric virus and the chimeric virus are useful for identifying cell
CC   lines capable of supporting the replication of these chimeric viruses, in
CC   screening for neutralizing antibodies to HCV of different genotypes, in
CC   the production of HCV-BVDV virions, for the development of inactivated or
CC   attenuated vaccines to prevent HCV-BVDV in a mammal, in studying the
CC   molecular properties of HCV indirectly in vitro, and in identifying
CC   inhibitors of viral enzyme activity which would be useful as antiviral
CC   agents. Formulations or compositions comprising the chimeric virions may
CC   be used to treat or prevent the signs and symptoms of HCV. The present
CC   sequence is encoded by a HCV clone, which is used to construct chimeric
CC   nucleic acids of the invention
XX
SQ   Sequence 3033 AA;
```

Query Match   21.3%;   Score 3277.5;   DB 4;   Length 3033;

Mon Oct 31 10:27:44 2005

us-10-009-002-2.rag

```
Best Local Similarity  30.3%;  Pred. No. 1.8e-244;
Matches  987;  Conservative  436;  Mismatches 1197;  Indels  641;  Gaps  117;

Db    903  AYLLKGALTRVPYFVRAHAL---LRMCTMARHLAGGRYV-QMALLALGRWTGTYIYDHLT   958
Qy    872  -HGDVLPNDFASKLPLQEP--FFPFEGKARVYRNEGRKLACGDTVDGLPVVARLGDLVPA   928
Db    959  PMSDWAASGLRDLAVAVEPIIFSPMEKKVIVWGAB--TAACGDILHGLPVSARLGREVLL  1016
Qy    929  GLAMPPD------GWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSY   983
Db   1017  G---PADGYTSKGWSLLAPITAYAQQTRGLLGTIVVSMTGRDKTEQAGEIQVLSTVTQSP  1073
Qy    984  MGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCSCGETK  1043
Db   1074  LGTSISGVLWTVYHGAGNKTLAGSRGPVTQMYSSAEGDLVGWPSPPGTKSLEPCTCGAVD  1133
Qy   1044  GYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGG-  1102
Db   1134  LYLVTRNADVIPARRRGDKRGALLSPRPLSTLKGSSGGPVLCPRGHAVGVFRAAVCSRGV  1193
Qy   1103  --SVSQIRVRPLVCAGHPQVTAHATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQ   1160
Db   1194  AKSIDFIPVETLDIVTRSPTFSDNS---TPPAVPQTYQVGVLHAPTGSGKSTKVPAYAA   1250
Qy   1161  EKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYLT-GACS  1219
Db   1251  QGYKVLVLNPSVAATLGFGAVLSKAHGINPNIRTGVRTVTTGAPITYSTYGKFLADGGCA  1310
Qy   1220  RN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANITEIQ  1278
Db   1311  GGAYDIIICDECHAVDSTTILGIGTVLDQAETAGVRLTVLATATPPGSVTTPHPNIEEVA  1370
Qy   1279  LTDEGTIPFHGKKIKEENLKKGRHLIFEATKGHCDELANELARKGITAVSYYRGCDISKI  1338
Db   1371  LGQEGEIPFYGRAIPLSYIKGGRHLIFCHSKKKCDELAAALRGMGLNSVAYYRGLDVSVI  1430
Qy   1339  P-EGDCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQ   1397
Db   1431  PTQGDVVVVATDALMTGYTGDFDSVIDCNVAVTQVVDFSLDPTFTITQIVPQDAVSRSQ   1490
Qy   1398  RRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAKAWYGLSSTRAQTILDTYRTQP   1457
Db   1491  RRGRTGRGRLGIYRYVSTGERASGMFDSVVLCECYDAGAAWYELTPSETTVRLRAYFNTP  1550
Qy   1458  GLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNVVLLTAAQLQLCHQYGYAAPNDAP  1516
Db   1551  GLPVCQDHLEFWEAVFTGLTHIDAHFLSQTKQSGENFAYLTAYQATVCAR----AKAPPP  1606
Qy   1517  RWQGARLGKKPCGVLWRLDGADAC------------PGPEP--                   1551
Db   1607  SWD----------VMWK----------CLTRLKPTLVGPTPLLYRLGSVTNEVTLTHPVTKYIA  1650
Qy   1552  --QMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEE  1609
Db   1651  TCMQADLEVMTS-TWVLAGGVLAAVAAYCLAT-GCVCIIGRLHINQ--RAVVAP--DKE   1703
Qy   1610  EI------VEBCASFIPL-EAMVAAIDKLKSTI----------------TTTSPFTLETALEKLN  1651
Db   1704  VLYEAFDEMEECASRAALIEEGQRIAEMLKSKIQGLLQQASKQAQDTQP-TVQASWPKVE  1762
Qy   1652  TFIGPHAATILAIIEYCCGLVTLPDNPFASCVFAFIAGITTPPLPHKIKMFLSLFGGAIAS  1711
Db   1763  QFWAKHMWNFISGIQYLAGLSTLPGNPAVASMMAFSAALTSPLSTSTTILLNILGGWLAS  1822
Qy   1712  KLTDARGALAFMMAGAGTALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPT   1768
Db   1823  QIAPPAGATGFVVSGLVGAAVG---SIGLGKVLVDILAGYGAGISGALVAFKIMSGEKPS  1879
Qy   1769  MDQLAGLVYSAFNPAAGVVGVLSACAMFALTTAGPD---HWPNRLLTMLARSNTVCNEY   1824
Db   1880  MEDVVNLLPGILSPGALVVGVI--CAILRRHVGPGEGAVQWMNRLIAFASRGNHVAPTH   1937
Qy   1825  FIATRDIRRKIIGILEASTPWSVISACIRWLHTPTEDDCGLI---AWGLEIWQYVCNFFV  1881
Db   1938  YVTESDASQRVTQLLGSLT----ITSLLRRLHNWITEDCPIPCGGSWLRDVWDWCTILT   1993


Qy      8  TSPVPAPRTRKNKQTQASYPVSIK-----------TSVERGQRAKRKVQRDARPR        51
Db      3  TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR   59
Qy     52  NYKIAGIHDGLQTLAQAALPAH----------------GWGRQDPRHKSRNL          87
Db     60  GRRQPIPKDRRSTGKSWGKPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPNDPRHRSRNV  119
Qy     88  GILLDYPLGWIGDVTVTHTPLVG-PLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLF  143
Db    120  GKVIDTLTCGFADLMGYIPVVGAPL-GGVARALAHGVRVLEDGVNFATGNLPGCSFSIF   177
Qy    144  VVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC------AD  197
Db    178  LLALLSCITTPVSAAEVKNISTG-YMVTNDCTNDSITWQLQAAVLHVPGCVPCEKVGNAS  236
Qy    198  ECWVPANPYIS--HPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV-  254
Db    237  QCWIPVSPNVAVQRPGALT---QGLRTHIDMVVMSATLCSALYVGDLCGGVMLAAQMFIV  293
Qy    255  ---RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTKLASQVPYAIAT  311
Db    294  SPQHHWFVQ------DCNCSIYPGITGHRMAWDMMMNWSPTATMILAYAMRVPEVIID    346
Qy    312  MFSSVHYLAVGALIYYASRGKKWYQLILALMLYIEATSGNPIRVPT-GCSIAE-------  362
Db    347  IISGAIWGVMFGLAYFSMQGAWAKVVVILLL----AAGVDARTHTVGGSAAQTTGRLTSL  402
Qy    363  ----------------------------FCSPLMIP----------CP------CHS-    375
Db    403  FDMGPRQKIQLVNTNGSWHINRTALNCNDSLHTGFIASLFYTHSFNSSGCPERMSACRSI  462
Qy    376  ------------YLSENVSE--------------------VICYSPKWTRPITLE       398
Db    463  EAFRVGWGALQYEDNVTNPEDMRPYCWHYPPRQCGVVSAKTVCGPVYCFTPS---PVVVG  519
Qy    399  YNNSISWYPYTIPGARGCMVKFKNNT-------WGC-----------------CRIR---N  432
Db    520  TTDRLGAPTYTW-GENETDVFLLNSTRPPLGSWFGCTWMNSSGYTKTCGAPPCRTRADFN  578
Qy    433  VPSYCTMGTDAVWNDTRNTYEACGVTPWLT----------------TAWHNGSALKLAILQ---YPG  480
Db    579  ASTDLLCPTDCPRKHPDTTYLKCGSGPWLTPRCLIDYPYRLWHYPCTVNYTIFKIRMYVG  638
Qy    481  SKEMFKPHNWMSGHLYPEGSDTPIVYFYDPVNSTLLP----PERWARLPGTPPVVRGSWL  536
Db    639  GVE----HRLTAACNFTRGDRCNL----EDRDRSQLSPLLHSTTEWAILPCS--------  682
Qy    537  QVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTKAVVLILLGLCGSKYL   596
Db    683  --------YSDLPALSTGLLHLHQNIVDVQFMYGLSPALTKYIVRWEWILLLFLLLADAR--  734
Qy    597  ILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLR  656
Db    735  VCACLWMLIL---------------------LGQAEAALEKLVILHAA-------S      762
Qy    657  YAALLGFVPMAAGLPLTFFVAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALLVGPWP   716
Db    763  AASCNGFLYF-----WYIKGRVVPLATYS----------------LTGLWS          799
Qy    717  LVALLTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQRW  776
Db    800  F--SLLLLALPQQAYAYDASVHGQIGAALLVMITLFTLTPGYKTLLSR------FLW--W  849
Qy    777  ------------------ENWFWNVTLRPER----FFLVLVCFPGATYDALVTFCVCHVALLCLTSS  821
Db    850  LCYLLTLGEAMVQEWAPPMQVRGGRDGIIWAVAIFYPGVVFDITKW------LLAVLGP   902
Qy    822  AASFFGTDSR-----VRAHRMLVRLGKC----HAWYSHYVLKFFLLVFGE-NGVFFYKHL-  871
```

```
Qy  2843 LVKYLAVIVFALGLIAVGLAI 2863
              :||::|| ||:|
Db  3014 L------LFGLLLLFVGVGL 3027
```

```
RESULT 9
AAB30729
ID   AAB30729 standard; protein; 3033 AA.
XX
AC   AAB30729;
XX
DT   02-APR-2001 (first entry)
XX
DE   Amino acid sequence of infectious Hepatitis C virus strain HC-J6CH.
XX
KW   HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine.
XX
OS   Hepatitis C virus.
XX
XX
PN   WO200075338-A2.
XX
PD   14-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015446.
XX
PR   04-JUN-1999; 99US-0137693P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Yanagi M, Bukh J, Emerson SU, Purcell RH;
XX
DR   WPI; 2001-061728/07.
DR   N-PSDB; AAC86644.
XX
PT   Nucleic acid molecule encoding human hepatitis C virus of genotype 2a for
PT   developing vaccines, for diagnosis of hepatitis C virus and in screening
PT   assays for identification of antiviral agents.
XX
PS   Disclosure; Page 88-99; 167pp; English.
XX
CC   The present sequence represents an amino acid sequence of infectious
CC   Hepatitis C virus (HCV) strain HC-J6CH genotype 2a. The HCV
CC   polynucleotide sequence is capable of expressing the virus when
CC   transfected into cells. The HCV protein is useful for assaying candidate
CC   antiviral agents for activity against HCV. Antibodies specific for HCV
CC   polypeptide are useful in prevention and treatment of diseases caused by
CC   HCV in animals, in particular humans. The HCV polypeptides serve as
CC   immunogens in the development of vaccines for preventing HCV in mammals
CC   or as antigens in diagnostic assays for detecting the presence of HCV in
CC   biological samples. The HCV polynucleotide is also useful for identifying
CC   cell lines capable of supporting the replication of HCV in vitro and to
CC   produce attenuated viral strains via passage in vitro or in vivo
XX
SQ   Sequence 3033 AA;

Query Match    21.3%;    Score 3277.5;    DB 4;    Length 3033;
Best Local Similarity   30.3%;    Pred. No. 1.8e-244;
Matches 987; Conservative 436; Mismatches 1197; Indels 641; Gaps 117;
```

```
Qy    8 TSPVAPRTRKNKQTQASYPVSIK-----------------TSVERGQRAKRKVQRDARPR 51
          :||:::|                                  ::||
Db    3 TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR 59

Qy   52 NYKIAGIHDGLQTLAQAALPAH------------------GWGRQDPRHKSRNL 87
Db   60 GRRQPIPKDRRSTGKSWGKPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPNDPRHRSRNV 119

Qy   88 GILLDYPLGWIGDVTHTPLVG-PLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLF 143
Db  120 GKVIDTLTCGFADLMGYIPVVGAPL--GGVARALAHGVRVLEDGVNFATGNLPGCSFSIF 177

Qy  144 VVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC-----AD 197
```

```
Qy 1882 ICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLY 1941
Db 1994 DFKNWLTS--KLFPKMPGLPFVSCQKGYKGVWAGTGIMTTRCPCGANISGNVRLGSMRI- 2050

Qy 1942 KGPRTCSNYWRGAVPVNARLCGSARPDT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVS 2000
Db 2051 TGPKTCMNIWQGTFPINCYTEQCVPKPAPNPKVAIWRVAASEYAEVTQHGSYHYITGLT 2110

Qy 2001 SPNVCFTQVP---PTLRAAVAVDGVQVQCYLGEPKTPWTTS-ACCYGPDGKGTVKLPFR 2056
Db 2111 TDNL---KVPCQLPSPEFFSWVDGVQIHRFAPTPKPFFRDEVSFCVGLNSFVVGSQLPCD 2167

Qy 2057 VDGHTPGVRMQLNLRD----ALETNDCNSTNNTPSDEAAVSA-LVFKQELRRT------- 2104
Db 2168 PEPDT-DVLMSM-LTDPSHITAETAARRLIARGSPPSEASSASQLSRATCTTHGKA 2225

Qy 2105 --NQLLEA-ISAGVDTTKLPA------------------PSI-EVVVRKRQFRART 2139
Db 2226 YDVDMVDANLFMGGDVTRIESGGSKVVVLDSLDPMVEERSDLEPSIPSEYMLPKKRF---- 2281

Qy 2140 GSLTLPP--PPRSVPGVSCP--ESLQRSD----------PLEGPSNLPPSPP------- 2177
Db 2282 -----PPALPAWARPDYNPPLVESWKRPDYQPATVAGCALPPPRKTPTPPPRRRRTVGLS 2336

Qy 2178 ------VLQLAM------PMPLLGAGECNPFTAIGC------AMTETGG--------- 2208
Db 2337 EDSIGDALQQLAIKSFGQPPP---SGDSGLSTCGAGAADSGSQTPPDELALSETGSISSMP 2393

Qy 2209 -----GPDDL---PSYPPKKEV-SEWSDE-SWSTATTASSYVTGPPYPKIRGKDST 2254
Db 2394 PLEGELGDPDDLPEQVEPQPPPQGGVAAPGSDSDGSWSTCSEEDDSVV---------- 2440

Qy 2255 QSAPAKRPTKKKLGKSEFSCSMSYTWTD-VISFKTASKVLSATRAITSGFLKQRSLVVVT 2313
Db 2441 -----------CCSMSYWTGALITPCSPEEKLPINPLSNSLLRVHNKVYCT 2482

Qy 2314 EPRDAELRKQKVTINRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSH 2373
Db 2483 TTKSASLRAKKVTFDRMQVLDSYDSVLKDIKLAASKVTARLLTMEEACQLITPPHSARSK 2542

Qy 2374 ITGLRGTDVR--SGAARKAVLDLQK-CVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKP 2430
Db 2543 Y-GFGAKEVRSLSGRAVNHIKSVWKDLLEDSETPI---PTTIMAKNEVFCVDPTKGGKKA 2598

Qy 2431 PRLISYPHLEMRCVEKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMWS--PDAVG 2487
Db 2599 ARLIVYPDLGVRVCEKMALYDITQKLPQAVMGASYGFQYSPAQRVEFLLKAWAEKKDDPMG 2658

Qy 2488 ATCDTVCFDSTITPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGY 2547
Db 2659 FSYDTRCFDSTVTERDIRTEESIYRACSLPEEAHTAIHSLTERLIYVGGPMFNSKGQTCGY 2718

Qy 2548 RRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMGNPRFLICGDDCTVIWKSAGADADKQAM 2607
Db 2719 RRCRASGVLTTSMGNTITCYVKALAACKAAGIIAPTMLVCGDDLVVISESQGTEEDERNL 2778

Qy 2608 RVFASWMKVMGAPQDCVPQPKYSLEELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCS 2667
Db 2779 RAFTEAMTRYSAPPGDPPRPEYDLELITSCSSNVSVALGPQGRRRYYLTRDPTTPIARAA 2838

Qy 2668 AEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPV 2727
Db 2839 WETVRHSPVNSWLGNIIQYAPTIWARMVLMTHFFSILMAQDTLDQNLNFEMYGAVYSVSP 2898

Qy 2728 EDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGG 2787
Db 2899 LDLPAIIERLHGLDAFSLHTYTPHELTTVASALRKLGAPPLRAWKSRARAVRASLISRGG 2958

Qy 2788 AHAKLARFLL-WHATSR----PLPDLDKTSVARYTTFNYCDVVSPEGDVFITPQRRLQKF 2842
Db 2959 RAAVCGRYLFNWAVKTKLKLTPLPEARLLDLSSWFT------VGAGGGDIYHSVSRARPRL 3013
```

```
Qy  1103 --SVSQIRVRPLVCAGHPQYTAHATLDTKPTVPNEYSVQILLAPTGSGKSTKLPLSYMQ  1160
Db  1194 AKSIDFPIVETLDIVTRSPTFSDNS---TPPAVPQTYQVGYLHAPTGSGKSTKVPVAYAA  1250

Qy  1161 EKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSYGMYLT-GACS   1219
Db  1251 QGYKVLVLNPSVAATLGFGAYLSKAHGINPNIRTGVRTVTTGAPITYSYTYGKFLADGGCA 1310

Qy  1220 RN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANITEIQ  1278
Db  1311 GGAYDIIICDECHAVDSTTILGTGTVLDQAETAGVRLTVLATATPPGSVTTPHNIEEVA   1370

Qy  1279 LTDEGTIPPHGKKIKEENLKKGRHLIFEATKKHCDELANELARKGITAVSYYRGCDISKI  1338
Db  1371 LGQEGEIPFYGRAIPLSYIKGGRHLIFCHSKKKKCDELAAALRGMGLNSVAYYRGLDVSVI 1430

Qy  1339 P-EGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFMGVRVCGVSAIVKGQ   1397
Db  1431 PTQGDVVVVATDALMTGYTGDFDSVIDCNVAVTQVVDFSLDPTFTITQIVPQDAVSRSQ   1490

Qy  1398 RRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQP  1457
Db  1491 RRGRTGRGRLGIYRYVSTGERASGMFDSVVLCECYDAGAAWYELTPSETTVRLRAYFNTP  1550

Qy  1458 GLPAIGANLDEWADLFS-MVNEPSFVNTAKRTADNYVLLTAAQLQLCHQYGYAAPNDAP   1516
Db  1551 GLPVCQDHLEFWEAVFTGLTHIDAHFLSQTKQSGENPAYLTAYQATVCAR----AKAPPP  1606

Qy  1517 RWQGARLGKKPCGVLWRLDGADAC------PGPEP------SEVTRY-              1551
Db  1607 SWD------VMWK------CLTRIKPTLVGPTPLLYRLGSVTNEVTLTHPVTKYIA       1650

Qy  1552 --QMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEE  1609
Db  1651 TCMQADLEVMTS-TWVLAGGVLAAVAAYCLAT-GCVCIIGRLHINQ---RAVVAP--DKE  1703

Qy  1610 EI------VEECASFIPL-EAMVAAIDKLKSTI------TTTSPFTLETALEKLN        1651
Db  1704 VLYEAFDEMEECASRAALIEEGQRIAEMLKSKIQGLLQQASKQAQDIQP-TVQASWPKVE  1762

Qy  1652 TFLGPHAATILAIIEYCCGLVTLPDNPFASCVPAFIAGITTPLPHKIKMFLSLFGGAIAS  1711
Db  1763 QFWAKHMWNFISSGIQYLAGLSTLPGNPAVASMMAFSAALTSPLSTSTTILLNLIGGWLAS 1822

Qy  1712 KLTDARGALAFMMAGAAGTALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPT  1768
Db  1823 QIAPPAGATGFVVSGLVGAAVG---SIGLGKVLVDILAGYGAGISGALVAFKIMSGEKPS  1879

Qy  1769 MDQLAGLVYSAFNPAAGVVGVLSACAMFALTTAGPD----HWPNRLLTMLARSNTVCNEY  1824
Db  1880 MEDVVNLLPGILSPGALVVGVI--CAAILRRHVGPGEGAVQWMNRLIAFASRGNHVAPTH  1937

Qy  1825 FIATRDIRRKIGILEASTPWSVISACIRWLHTPTEDDCGLI----AWGLEIWQYVCNFFV  1881
Db  1938 YVTBSDASQRVTQLLGSLT----ITSLLRRLHNWITEDCPIPCGGSWLRDVWDWVCTLLT  1993

Qy  1882 ICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLY  1941
Db  1994 DFKNWLTS--KLFPKMPGLPFVSCQKGYKGVWAGTGIMTTRCPCGANISGNVRLGSMRI-  2050

Qy  1942 KGPRTCSNYWRGAVPVNARLCGSARPDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVS  2000
Db  2051 TGPKTCMNIWQGTFPINCYTEGQCVPKPAPNFKVAIMRVAASEYAEVTQHGSYHYITGLT  2110

Qy  2001 SPNVCFTQVP---PTLRAAVAVDGVQVQCYLGEPKTPWTTS-ACCYGPDGKGKTVKLPFR  2056
Db  2111 TDNL---KVPCQLPSPEFFSWVDGVQIHRFAPTPKPFFRDEVSFCVGLNSFVVGSQLPCD  2167

Qy  2057 VDGHTPGVRMQLNLRD------ALETNDCNSTNNTPSDEAAVSA-LVFKQELRRT------ 2104
Db  2168 PEPDT-DVLMSM-LTDPSHITAETAARRLARGSPPSEASSASQLSAPSLRATCTTHGKA   2225
```

```
Db   178 LLALLSCITTPVSAAEVKNISTG-YMVTNDCTNDSITWQLQAAVLHVPGCVPCEKVGNAS  236
Qy   198 ECWVPANPYIS--HPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV-  254

Db   237 QCWIPVSPNVAVQRPGALT--QGLRTHIDMVVMSATLCSALYVGDLCGGVMLAAQMFIV   293
Qy   255 ---RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTKLASQVPYAIAT  311

Db   294 SPQHHWFVQ------DCNCSIYPGTITGHRMAWDMMMWSPTATMILAYAMRVPEVIID    346
Qy   312 MFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPT-GCSIAE------    362

Db   347 IISGAHWGVMFGLAYFSMQGAWAKVVVILLL---AAGVDARTHTVGGSAAQTTGRLTSL   402
Qy   363 ----------------------FCSPLMIP-------CP-----CHS-            375

Db   403 FDMGPRQKIQLVNTNGSWHINRTALNCNDSLHTGFIASLFYTHSFNSSGCPERMSACRSI  462
Qy   376 ----------YLSENVSB-------------VICYSPKWTRPITLE                398

Db   463 EAFRVGWGALQYEDNVTNPEDMRPYCWHYPPRQCGVVSAKTVCGPVYCFTPS---PVVVG  519
Qy   399 YNNSISWYPYTIPGARGCMVKFKNNT------WGC------CRIR---N             432

Db   520 TTDRLGAPTYIW-GENETDVFLLNSTRPPLGSWFGCTWMNSSGYTKTCGAPPCRTRADFN  578
Qy   433 VPSYCTMGTDAVWNDTRNTYEACGVTPWLT---------TAWHNGSALKLAILQ--YPG   480

Db   579 ASTDLLCPTDCFRKHPDTTYLKCCGSGGPWLTPRCLIDYPYRLMHYPCTVNYTIFKIRMYVG 638
Qy   481 SKEMFKPHNWMSGHLYFEGSDTPIVVFYDVPVNSTLLP----PERWARLPGTPPVVRGSWL 536

Db   639 GVE----HRLTAACNFTRGDRCNL--EDRDRSQLSPLLHSTTEWAILPCS------      682
Qy   537 QVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYL  596

Db   683 ------YSDLPALSTGLLHLHQNIVDVQFMYGLSPALTKYIVRWEWVILLFLLADAR-   734
Qy   597 ILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLR  656

Db   735 VCACLWMLIL---------LGQAEAALEKLVILHAA------S                   762
Qy   657 YAALLGFVPMAAGLPLTFFVAAAAQPDYDKWVRLLVAGLVLWAGRNRGHRIALLVGPWP   716

Db   763 AASCNGFLYF-----VIFFVAA-------WYIKGRVVPLATYS-------LTGLWS      799
Qy   717 LVALLTILHLVTPASAFDTEIGGLLTIPPVALVVMSRFGFFAHILLPRCALVNSYLWQRW  776

Db   800 F--SLLLLALPQQAYAYDASVHGQIGAALLVMITLFTLTPGYKTLLSR-----FLW--W   849
Qy   777 -------ENWFWNVTLRPER-----FFLVLVCFPGATYDALVTFCVCHVALLCLTSS     821

Db   850 LCYLLTIGEAMVQEWAPPMQVRGGRDGIIWAVAIFYPGVVFDITKW----LLAVLGP    902
Qy   822 AASFFGTDSR-----VRAHRMLVRLGKC----HAWYSHYVLKFFLLVFGE-NGVFFYKHL- 871

Db   903 AYLLKGALTRVPYFVRAHAL--LRMCTMARHLAGGRYV-QMALLALGRWTGYIYDHLT   958
Qy   872 -HGDVLPNDFASKLPLQEP--FFPFEGKARVYRNEGRRLACGDTVDGLPVARLGDLVFA   928

Db   959 PMSDWAASGLRDLAVAVEPIIFSPMEKKVIVWGAE--TAACGDILHGLPVSARLGREVLL  1016
Qy   929 GLAMPPD-------GWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIPRLGSLATSY 983

Db  1017 G----PADGYTSKGWSLLAPITAYAQQTRGLLGTIVVSMTGRDKTEQAGEIQVLSTVTQSF 1073
Qy   984 MGFVCDNVLVLTYAHHGSKGRRLAHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCSSCGETK 1043

Db  1074 LGTSISGVLWTVYHGAGNKTLAGSRGPVTQMYSSAEGDLVGWPSPPGTKSLEPCTCGAVD  1133
Qy  1044 GYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGG-  1102

Db  1134 LYLVTRNADVIPARRRGDKRGALLSPRPLSTLKGSSGGPVLCPRGHAVGVFRAAVCSRGV  1193
```

```
KW   Non A non B hepatitis virus; plasma; degenerate.
XX
OS   Non-A.
OS   non-B hepatitis virus.
XX
PN   EP532167-A2.
XX
PD   17-MAR-1993.
XX
PF   30-JUL-1992;    92EP-00306952.
XX
PR   09-AUG-1991;    91JP-00287402.
PR   05-DEC-1991;    91JP-00360441.
XX
PA   (IMMO ) IMMUNO JAPAN INC.
XX
PI   Okamoto H, Nakamura T;
XX
DR   WPI; 1993-087166/11.
DR   P-PSDB; AAQ38221.
XX
PT   Polynucleotide(s), polypeptide(s) and antibodies of NANBH virus - useful
PT   for detecting NANBH, as a vaccine and for screening blood samples.
XX
PS   Claim 8; Page 65-79; 93pp; English.
XX
CC   RNA was isolated from the plasma of human patients positive for NANBH
CC   virus (strain HC-J8) and was subjected to reverse transcription to
CC   produce cDNA. The resulting cDNA was amplified by PCR, and nucleic acid
CC   sequences determined by analysis of clones obtd. by PCR amplification (42
CC   clones in total). The NANBH HC-J8 genome was found to contain a
CC   degenerate open reading frame encoding polypeptide precursors of 3033
CC   amino acid residues. See also AAR33538 and AAR33214. (Updated on 25-MAR-
CC   2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 3033 AA;
```

```
Query Match             21.3%;  Score 3268.5;  DB 2;  Length 3033;
Best Local Similarity   30.0%;  Pred. No. 9.1e-244;
Matches  983;  Conservative  433;  Mismatches 1185;  Indels  679;  Gaps  106;

Qy         8  TSPVPAPRTRKNKQTQASYPVSIK-----------------TSVERGQRAKRKVQRDARPR   51
              |::|| ||||| | | :| | :||                  |:||| : |:| ||| ||
Db         3  TNPKP---QRKYKRNTNRRPQDVKFPGGGQIVGVYLLPRRGPGVRATRKTSERSQPR   59

Qy        52  NYKIAGIHDGLQTLAQAALPAH-----------------GWGRQDPRHKSRNL   87
              | :||: || :: | || ||                    || || |||||
Db        60  GRRQPIPKDRRSTGKSWGKPGYPWPLYGNEGCGWAGWLLSPRGSRPTWGPTDPHRSRNL  119

Qy        88  GILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG--------WFGVHLFV  144
              | :| ||| || :||  || | ||| |  | |  ::|| | ||: |           |||: ||
Db       120  GRVIDTITCGFADLMGYIPVGGAPV-GGVARALAHGVRVLEDGINYATCNLPGCSFSIFL  178

Qy       145  VCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVICADE------  198
              |||:  ||  | | | ::| :| |:|||  :||  |    ||||  |   || |
Db       179  LALLSCVTVPVSAVEVRN-ISSSYYATNDCSNNSITWQLTDAVLHLPGCVPCENDNGTLH  237

Qy       199  CWVPANPYISHPSNWTGTDSFLADHIDFVMGALLVTCDALDIGELCGACVLVGDWLV----  254
              ||  | |:||    ||  |: |  ||| | :: |::: || ::  |: | | | | :
Db       238  CWIQVTPNVAVKHRGALTRS-LRSHVDMIVMAATACSALYVGDVCGAVMILSQAFMVSPQ  296

Qy       255  RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTKLASQVPYAIATMFS  314
              || :  ::|  |   ||:|::|  ||   |:|:||  |||:| |: |||  | |:| |
Db       297  RH------NFTQECNCSIYQGHITGHRMAWDMMLSWSPTLTMILAYARVPELVLEIIF  349

Qy       315  SVHYLAVGALIYYASRGKWYQLLLALMLY-IEAT----SGNPIRVPTGCSIAEFCSPLMI  369
              |: || | | :|| ||| || : || |   :|       | || || |  :| | ||
Db       350  GGHWGVVFGLAYFSMQGAWAKVIAILLLVAGVDATTYSSGQ----EAGRTVAGFAGLFTT  405

Qy       370  PCPCHSYLSE------NVSEVIC------------------YSPKWT------------RPI---  395
              | || ||| |      |: |  |                  |||||              ||:
Db       406  GAKQNLYLINTNGSWHINRSALNCNDSLQTGFLASLFYTHKFNSSGCPERLSSCRGLDDF  465
```

```
RESULT 10
AAR33539
ID   AAR33539 standard; protein; 3033 AA.
XX
AC   AAR33539;
XX
DT   27-AUG-2003  (revised)
DT   25-MAR-2003  (revised)
DT   01-JUL-1993  (first entry)
XX
DE   NANBH virus strain HC-J8 protein.
XX
```

```
Qy      2105  --NQLLEA-ISAGVDTTKLPA----------------PSI-EEVVRKRQFRART  2139
                ::|:| | | | || |||                  ||| |:||:|:|| ||
Db      2226  YDVDMVDANLFMGGDVTRIESGSKVVLDSLDPMVEERSDLEPSIPSEYMLPKKRF-----  2281

Qy      2140  GSLTLPP--PPRSVPGVSCP--ESLQRSD----------PLEGPSNLPPSP---------  2177
              |  | ||  || :| | | |  ||| :|           || ||:|||||
Db      2282  ----PPALPAWARPDYNPPLVESWKRPDYQPATVAGCALPPPRKTPTPPPRRRRTVGLS  2336

Qy      2178  -------VLQLAM------PMPLLGAGECNPFTAIGC-----------AMTETGG-----  2208
                      ||| ||      |||||    | | | ||            || | ||
Db      2337  EDSIGDALQQLAIKSFQPPP---SGDSGLSTGAGAADSGSQTPPDELALSETGSISSMP  2393

Qy      2209  ------GPDDL-------PSYPPKKEV-SEWSDE-SWSTATTASSYVTGPPYPKIRGKDST  2254
                    || ||       ||| | |: |: | |  |: | | || :|| ||| || ||
Db      2394  PLEGELGDPDLEPQVEPQPPQGGVAAPGSDSGSWSTCSEEDDSVV----------------  2440

Qy      2255  QSAPAKRPTKKKLGKSEFSCSMSYTWTD-VISPKTASKVLSATRAITSGFLKQRSLVYVT  2313
              | : |:|| | || : | ||||  ::| | | | | :| || || |:| || ::::| |
Db      2441  --------------------CCSMSYSWTGALITPCSPEEKLPINPLSNSLLRYHNKVCT  2482

Qy      2314  EPRDAELRKQKVTINRQPLFPPSYHKQVRLAKEKASKVGVWMDYDEVAAHTPSKSAKSH  2373
              |||: || | ||:||:||:||| |||:|||| | || | ||  ||: | || ||||
Db      2483  TTKSASLRAKKVTFDRMQVLDSYYDSVLKDIKLAASKVTARLLTMEERACQLTPPHSARSK  2542

Qy      2374  ITGLRGTDVR--SGAARKAVLDIQK-CVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKP  2430
              || | ||| |  |||| || |::|  ||: |:||||||| :||||:|| |||| || ||
Db      2543  Y-GFGAKEVRSLSGRAVNHIKSVWKDLLEDSETPI---PTTIMAKNEVFCVDPTKGGKKA  2598

Qy      2431  PRLISYPHLEMRCVEKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMWS--PDAVG  2487
              || || ||| || | :|: |::|  | | | |::|   ||||  |: | :|    || |
Db      2599  ARLIVYPDLGVRVCEKMALYDITQKLPQAVMGASYGFQYSPAQRVEFLLKAWAEKKDPMG  2658

Qy      2488  ATCDTVCFDSTITPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGY  2547
              || ||  |||: || || || || ||| |  | |||  :| ||:: || | |::|::|
Db      2659  FSYDTRCFDSTVTERDIRTEESIYRACSLPEEAHTAIHSLTERLYVGGPMFNSKGQTCGY  2718

Qy      2548  RRCRSSGVYTTSSNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAM  2607
              |||:| |||| ||  :|| |:|  | |  || |:| ||||||| |:|||| |   | |
Db      2719  RRCRASGVLTTSMGNTITCYVKALAACKAAGIIAPTMLVCGDDLVVISESQGTEEDERNL  2778

Qy      2608  RVFASWNKMKVMGAPQDCVPQPKYSLEELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCS  2667
              |  |: | | | ||| | |  |  :|  |:  ||| : |:| | | :|||||  || ||
Db      2779  RAFTEAMTRYSAPPGDPRPEYDLELITSCSSNVSVALGPQGRRRYYLTRDPTTPIARAA  2838

Qy      2668  AEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPV  2727
              | || | :| | ||| |:||||||  |::| || |  |: | | :| |||| || | ||
Db      2839  WETVRHSPVNSWLGNIIQYAPTIWARMVLMTHFFSILMAQDTLDQNLNFEMYGAVYSVSP  2898

Qy      2728  EDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGG  2787
              || | :|| || | | | | | | ||||  ||  | || |||: || | |:|| |||
Db      2899  LDLPAIIERLHGLDAFSLHTYTPHELTRVASALRKLGAPPLRAWKSRARAVRASLISRGG  2958

Qy      2788  AHAKLARFLL-WHATSR-----PLPDLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKF  2842
              |  |: | |  |:| ||      |:| |   | | | | ||||  |  |  | :||:|:
Db      2959  RRAAVCGRFLFNNAVKTKLKLTPLPEARLLDLSSWFT-----VGAGGGDIYHSVSRARPRL  3013

Qy      2843  LVKYLAVIVFALGLIAVGLAI  2863
              |   |  :|| | |:||| | :
Db      3014  L------LFGLLLLFGVGL  3027
```

```
Qy   396 ----TLEYNNSIS-----------------W-YP------------YTIPDSARGCM----------- 417
             ||:||||                     |:|                ||||||
Db  1418 AVAYYRGLDVSVIPTQGDVVVATDALMTGYTGDFDSVIDCNVAVSQIVDFSLDPTFTIT 1477

Qy   466 RIGWGTLEYENTVTNDGDMRPYCWHYPPRPCGIVPARTVCGPVYCFTPSPVVVGTTDKQG 525
             ||:||||||||                                      |||||||
Db  1385 VRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFPDAAKAWYGLSST 1444

Qy   418 --------------VKFKNT--------WGC-----------CRIR----NVPSYC----- 437
                        :|||:            |||            |||:    |||||
Db  1478 TQTVPQDAVSRSQRRGRTGRGLGVYRYVSSGERPSGMFDSVVLCECYDAGAAWYELTPA 1537

Qy   526 VPTYTWGENETDVFLLNSTRPPRGAWFGCTWMNGTGFTKTCGAPPCRIRKDYNSTIDLLC 585
             ||||:    ::|  ||||   |  :||| :|||  |   :|||  |||  |||||||||||
Db  1538 ETTVRLRAYFNTPGLPVCQDHLEFWRAVFTGLTHIDAHFLSQTKQGGENFAYLTAYQATV 1597

Qy   438 TMGTDAVWNDTRNTYEACGVTPWLT-----------TAWHNGSALKLAILQ----YPGGSKEMF 485
             ||| | :||| |||:|  :||| ||           ||  ||:||| ||||    |||||
Db  1586 P---TDCFRKHPDATYLKCGAGPWLTPRCLVDYPYRLMWHYPCTVNFTIFKARMYVGGVE-- 641

Qy   486 KPHNWMSGHLYF----------EGSDTPIVFYFYDPVNSTLLPPERWARLPGTPPVVRGS 534
             |||:                  |||||                :|:|
Db  1598 CAR------AKAPPSWD------------VMWK------CLTRLKFTLTGPTPLLYRLGAVTN 1637

Qy   535 WLQVPQGFYSDVKDLATGLIITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSK 594
             |||||                                  :|: |
Db  1546 ------SEVTRY----QMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITS- 1595

Qy   642 --HRFSAACNFTRGDRCRLEDRDRGQQSPLLH---------ST----TEWAVLPCS------- 682
                 |:||||:                         :|    |:|||
Db  1638 EVTLTHPVTKYIATCMQADLEIMTS-SWVLAGGVLAAVAAYCLAT------GCISIGR 1689

Qy   595 YLILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCR 654
               :|||        :|::::|                    ||||:   :|::|:
Db  1596 --VPTGATVAPVVDEE---------EIVEECASFIPL-EAMVAAIDKLKSTI---------TTTSP 1640

Qy   735 --------------ILGQAEAALEKLII----------- 757
                          :||:| :|:||
Db  1690 LHLNDRVVVAP--DKEILYEAFDEMBECASKAALIEEGQRMAEMLKSKIQGLLQQATRQA 1747

Qy   655 LRYAALLGFVPEMAAGLPLTFVAAAAAQPDYDWWVRLLVAGLVLWLAGRNRGHRIALLVGP 714
                |  |:|                       |::|| :|  :|| ||:|
Db  1641 FTLETALE----KLNTFLGPHAATILAIIEYCCGLVTLPDNPPASCVPAFIAGITTPLPH 1696

Qy   758 ------------WYLKGRVVPVATYS--------------VLGL 797
                        :|:||:|||                  ||||
Db  1748 QDIQPAIQSSWPKLEQFWAKHMWNFISGIQYLAGLSTLPGNPAVASMMAFSAALTSPLPT 1807

Qy   715 WPLVALLTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLW-- 773
                |:|::|::|:|              ||:|    |:|:|| ||||||  |:||:|
Db  1697 KIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVG---FVFDMLGGYAAASS 1753

Qy   798 WSF--LLLVLALPQQAYALDAAEQGELGLAILVIISIFTLTPAYKILLSRSVWWLSYMLV 855
             ||:   :::|:|                    ||:|  :||  :|:|::|:|::
Db  1808 STTILLNIMGGWLASQIAPPAGATGFVVSGLVGAAVG--SIGLGKILVDDVLAGYGAGIS 1864

Qy   774 -----QRW---------ENWFW-NVTLRPERFFLV----LVCFPGATYDALVTFCVCHV 813
               :|:            :|:|  ::| ||:||:     ||| ||::|||:| :| ||:
Db  1754 TACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVGVLSACAMPFALTTAGPD--HWPNRLL 1811

Qy   856 LAEAQIQQWVPPLEVRGGRDGIIWVAVILHPRLVFEVTFKWLLAILGPAY---------- 904
                 |::|                 |:|  :|:
Db  1865 GALVAFKIMSGEKPTVEDVVNLLPAILSPGALVVGVICAAILRRHVGQGEGAVQWMNRLI 1924

Qy   814 ALLCLTSSAASFFGTDSRVRAHRML--VRLGKCHAWYSHYVLKFFLLVFGENGVFFYKHL 871
             | :|:|              :|:|:                |  :|::|:
Db  1812 TMLARSNTVCNEYFIATRDIRRKILGLLEASTPWSVISACIRWLHTPEDDCGLI----AW 1868

Qy   905 -LLKASLLRIPYF-----VRAHALLRVCTLVK-HLAGARYIQMLLITIGRWTGTYLYDHL 957
                :|:|:|:      |:|:|:| |::|   ||:|||:::|:|::|  :|
Db  1925 APASRGNHVAPTHYVVESDASQRVTQVLSSLT----ITSLLRRLHAWITEDCPVPCSGSW 1980

Qy   872 HGDVLPNDFASKLPLQ----------EP---FFPPEGKARVYRNEGRRLACGDTVDG 915
             ||:  |:||         :|   ||||:||:| |::|:|
Db  1869 GLEIWQYVCNPFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAE 1928

Qy   958 S---------PLSTWAAQGLRDLAIAVEPVVFSPMEKKVIVWGAE--TVACGDILHG 1003
                      :|||:                     ::|:|||:   || ||:|
Db  1981 LQDIWDWVCSILTDPKNWLSS--KLLPKMPGIPFISCQKGYKGVWAGTGVMTRCPCGAN 2038

Qy   916 LPVVARLGDLVFAGLAMPPD------GWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWT 970
             ||:: ::|              |:||  :|: |||:|  ||| |::|  |:|
Db  1929 LIPFSVENGFAKLYKGPRTCSNTWRGAVPVNARLCGSARP-DPTDWTSLLVVNYGVRDYCKY 1987

Qy  1004 LPVSARLGREVLLG---PADGYTSKGWKLLAPITAVTQQTRGLLGAIVVSLTGRDKNEQA 1060
             ||:|:||:|:|||      |::|||::||:|| |||  |:||||||::|:|| ||:|:|
Db  2039 ISGHVRMGTMKI-TGPKTCLNLWQGTFPINCYTEGPCVPKPPPNYKTAIWRVAASEVVEV 2097

Qy   971 GTIPRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDIYQPPCG 1030
             ||::|| |  :|||                |:|:||||::|  :|  |:|:|:::|||
Db  1988 EKMGDHIFVTAVSSPNVCFTQVP---PTLRAAVAVDGVQVQCYLGEPKTPWTTSACCYGP 2044

Qy  1061 GQVQVLSSVTQTFLGTSISGVLWTVHGAGNKTLAGPKGPVTQMYTSAEGDLVGWPSPPG 1120
             ||:||| |:|:|                 ::||||:|  :|::| |:  |:| |::|
Db  2098 TQHGSFSVTGLTSDNL---KVPCQVPAPEFFSWVDGVQIHRPAPVP----------GP 2143

Qy  1031 AGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHV 1090
             |||:|||||||:||||:|||||||::|||||:||||||||:|:|| |||||||||||||
Db  2045 DGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSD---EAAVSALVFKQEL 2101

Qy  1121 TKSLDPCTCGAVDLYLVTRNADVIPVRRKDDRRGALLSPRPLSTLKGSSGGPVLCSRGHA 1180
             |:| |||:|||::|:||||||:|||:|:||||:||||:|||:|||||| |||||||||||
Db  2144 FFRD---EVTFTVGLNSFVVGSQLPCDPEPDTEVLASMLTDPSHITAEAAARRLARGSPP 2200

Qy  1091 IGMFTA--ARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILLIAPTG 1147
             :|:|:|    |:||||:||||||||||||||||||||||||:||||||:|||||||||||
Db  2102 RRTNQLLEAISA------------------------GVDTTKLPA------------- 2122

Qy  1181 VGLFRAAVCARGVAKSIDFIPVESLDVATRTPSFSDNS----TPPAVPQSYQVGYLHAPTG 1237
             |||:|||||||||||::||:||||||||||:|:|||||    |||||||| |:||||||||
Db  2201 SQASSSASQLSAPSLKATCTTHKTAYDCDMVDANLFMGGDVTRIESDSKVIVLDSLDSMT 2260

Qy  1148 SGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTY 1207
             |||||||||||::||||||||||||||||||||||||||:||||||||||:||:|||:||
Db  2123 --------PSI-BEVVVRKRQFRARTGSLTLPPPRSVPGVSCCPESLQRSDPLEGPSNLPP 2174

Qy  1238 SGKSTKVPAAYASQGVKVLVLNPSVAATLGFGAYMSKAHGINPNIRTGVRTVTTGDSITY 1297
             ||||||||||||:|                  ||    :|||||:|||||:|||
Db  2261 EVEDDREPSVPSEYLIKRRKFPP----------ALPPWAR---------------PD 2292

Qy  1208 STYGMYLT--GACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPG 1265
             ||||:||:   | |:||:|||||||  |:||||||:|  |||:||||:||||||||||||
Db  2175 SPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKKE------------- 2221

Qy  1298 STYGKFIADGGKAAGAYDIIIICDECHSVDATTILGIGTVLDQAETAGVRLVVLATATPPG 1357
             |||| :|||| | |  :||:|||||| |:||||:|||:||||  :|||:||:|:|||||||
Db  2293 YNPVLIETWKRP------GYEPPTVLGCALPPT----PQTPVPPRRRAKVLTQDNVEGV 2343

Qy  1266 VIPTPHANITEIQLTDEGTIPHGKKIKEENLKKGRHLIFEATKKHCDELANELARKGIT 1325
              :||:  : |:::|          |:||   |:| |||  ||: ||||
Db  2222 -------------------------------VSEWSDESWSTATTASSVTGPPYPKIRGK 2251

Qy  1358 TVTTPHSNIEBVALGHEGEIPFYGKAIPLAFIKGGRHLIFCHSKKKCDELAAALRGMGVN 1417
             ||    |::|:|| |:|||:||||::| ||: ||||||||| |||:||||| |||:|:||
Db  2344 LREMADKVLSPLQDNNDSGHSTGADTGGDIVQQPSDE--TAASEAGLSSMPPLEGEPGD 2401

Qy  1326 AVSYYRGCDISKIP-EGDCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMG 1384
             ||:||||  | |   |||||||||||:||||||||||:|||||:||||||||||||||
Db  2252 DSTQ------SAPAKR-----------PTKKKLGKSEFSCSMSYTWTD-VISFKTASKV 2292
```

PT   polypeptide useful in gene therapy to treat hepatitis C.
XX
PS   Claim 1; Page 25-33; 36pp; Japanese.
XX
CC   The invention relates to a human polypeptide related to hepatitis C virus
CC   (HCV), and the polynucleotide encoding it. The polypeptide can be used
CC   for the development of gene therapy on fulminant hepatitis C. This
CC   sequence represents a human HCV-related polypeptide
XX
SQ   Sequence 3033 AA;

Query Match   21.2%; Score 3265.5; DB 5; Length 3033;
Best Local Similarity   29.8%; Pred. No. 1.6e-243;
Matches 964; Conservative 453; Mismatches 1232; Indels 583; Gaps 101;

Qy   8 TSPVPAPRTRKN--KQTQASYP-------------------VSIKTSVERGQRAKRKVQR 46
             :|:: ||::   .|::::   ::: :: :: .:::|:
Db   3 TNPKPQRKTKRNTNRRPEDVKFPGGGQIVGGVVLLPRRGPRLGVRTTRKTSERSQPRGRR 62

Qy   47 DARPRNYKIAGI-------------HDGLQTLAQAALP--AHGWGRQDPRHKSRNLGIL 90
             :|:|::: :.:   ::|:::::::: : ::::|:.|::::::::|::|
Db   63 QPIPKDRRSTGKAWGKPGRPWPLYGNEGLGWAGWLLSPRGSRPSWGPTDPRHRSRNVGKV 122

Qy   91 LDYPLGWIGDVTHTPLVGPLVAGAVVRPVCQIVRLEDGVNWATG--WFGVHLFVVCL 147
             :::|:|.|:|||.||:|||:||:|.|:|||:|::|::|:|:|:| |::|:||.||
Db   123 IDTLTCGFADLMGYIPVVGAPLSGA-ARAVAHGVRVLEDGVNYATGNLPGFPFSIFLLAL 181

Qy   148 LS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC----ADECWV 201
             :: :|.|.:||:|||:|||::.|::|:|::|:::::::::.|:.||::: ||:|.|
Db   182 LSCITVPVSAAQVKN-TSSSYMVTNDCSNDSITWQLEAAVLHVPGCVPCERVGNTSRCWV 240

Qy   202 PANP--YISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV----R 255
             .:: |::|.:||:||:.|:|::|||:|.|:|||||:|||:|||||:|.||:|: |
Db   241 PVSPNMAVRQPGALT--QGLRTHIDMVVMSATFCSALLYVGDLCGGVMLAAQVPIVSPQY 297

Qy   256 HWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTKLLASQVPYAIATMFSS 315
             || ::|::|:||:|.|::||.||:|:.||||::|:|:::||::.||||::|||:::||:::
Db   298 HWFVQ------ECNCSIYPGTITGHRMAWDMMNWSPTATMLAYVMRVPEVIIDIVSG 350

Qy   316 VHYLAVGALIYYASRGKWYQLLLALML-------------YIEAT---------- 342
             |:|:||:|::|:|:|:|::|:::|:|:   ::.|:
Db   351 AHWGVMFGLAYFSMQGAWAKVIVILLLAAGVDAGTTTVGGAVARSTNVIAGVFSHGPQQN 410

Qy   343 --SGNPIRVPTGCSIAEF---- 363
                                                       :||:|.||.||::|:|
Db   411 IQLINTNGSWHINRTALNCNDSLNTGFLAALFYTNRFNSSGCPGRLSACRNIEAFRIGWG 470

Qy   364 --------------CPCHSYLSENV-SEVICYSPK---------------- 390
                           |:|.|:|::| :|:||:|:|:
Db   471 TLQYEDNVTNPEDMRPYCWHYPPKPCGVPARSVCGPVYCFTPSPVVVGTTDRRGVPTYT 530

Qy   391 W--------------TRPITLEYNNSISWPYTIPGARGCMVKFKNNTWGC--CRIR---NV 433
             | .||::|:|::|:||||:|||:|||::|:|:|||.| |::| ||
Db   531 WGENETDVFLLNSTRP-----PQGSWFGCTWMNSTGF----TKTCGAPPCRTRADFNA 579

Qy   434 PSYCTMGTDAVWNDTRNTYEACGVTPWLT-----------TAWHNGSALKLAILQ--YPGS 481
             |::|::|:|:|:|:|:|:|::||:|||:| ||||:||:||:|::| |||:
Db   580 STDLLCPTDCFRKHPDATYIKCGSGPWLTPKCLVHYPYRLWHYPCTVNFTIKIRMYVGG 639

Qy   482 KEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLP-----PERWARLPGTPPVVRGSWLQ 537
             |:|:|:||||:|:||:||::||::|:||||::|::| |::||::|:||||||:|||:
Db   640 VE----HRLTAACNFTRGDRCDL---EDRDRSQLSPLLHSTTEWAILPCT---------- 682

Qy   538 VPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLI 597
             ||::|:|:|::|||.||::|::|::||:|:|::|||:::|||:|:|:|:|||.|||:|:|:
Db   683 -----YSDLPALSTGLLHLHQNIVDVDVQYMYGLSPAITKYVVRWEWVVLFLLLADAR--V 735

Qy   598 LAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYLRCR-LR 656
             ||:|||:|:|::|:|:|||:|:||:|||:|||||||:||||:|||::|:::|||:||| |:
Db   736 CACL-WMLILLGQAEAALEKLV------------------VLHAASAA------ 770

Qy   657 YAALLGFVPMAAGLPLIFFVAAAAAQPDYDWMVRLLVAGLVLWAGRNRGHRIALLVGPWP 716
             |.|:||||:|||||||:||:|||:|:||::||:.|:|:|::||||:|:|:|::|:||||

Mon Oct 31 10:27:44 2005   us-10-009-002-2.rag

```
Qy  1726  GAAGTALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNP  1782
Db  1837  GLVGAAVG---SIGLGKVLVDILAGYGAGISGALVAFKIMSGEKPSMEDVINLLPGILSP  1893

Qy  1783  AAGVGVLSACAMFALTTAGPD---HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGI    1838
Db  1894  GALVVGVI--CAAILRRHVGPGEGAVQWMNRLLIAFASRGNHVAPTHYVTESDASQRVTQL  1951

Qy  1839  LEASTPWSVISACIRWLHTPTEDDCGLI---AWGLEIWQYVCNFFVICFNVLKAGVQSMV   1895
Db  1952  LGSLT----ITSLLRRLHNWITEDCPIPCSGSWLRDVWDWVCTILTDFKNWLTS--KLFP   2005

Qy  1896  NIPGCPYYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAV   1955
Db  2006  KLPGLPFISCQKGYKGVWAGTGIMTTRCPCGANISGNVRLGSMRI-TGPKTCMNTWQGTF   2064

Qy  1956  PVNARLCGSARP-DPTDWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLR   2014
Db  2065  PINCYTEGQCAPKPPTNYKTAIWRVAASEYAEVTQHGSYSYYTGLTTDNLKIPCQLFSPE   2124

Qy  2015  AAVAVDGVQVQCYLGEPKTPWTTS-ACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDA   2073
Db  2125  FFSWVDGVQIHRFAPTPKPFFRDEVSFCVGLNSYAVGSQLPCEPEPDADVLRSML------   2179

Qy  2074  LETNDCNSTNNTPSDEAAVSALVFKQELRRTNQLLEAISAGVDTTKLPAPSIEEVVVRKR   2133
Db  2180  ------------TDPPHITAETAARRLARGSPPSEASSS---VSQLSAPSLRATCTT--    2221

Qy  2134  QFRART-------GSLTLPPPRSVPGVSCPESLQRSDPLEGPSNLPPSPPV--          2178
Db  2222  --HSNTYDVDMVDANLLMEGGVAQTEPESRVPVLDFLEPM----AEEESDLEPSIFSEC    2274

Qy  2179  -------PLLGA---GBCNPFTAIGCAMTETGGGPDDLPSYPPKKE                2221
Db  2275  MLPRSGFPRALPAWARPDYNPPLVESWRRPDYQPPTVAGCALPP---PKKAPTPPPRRR    2330

Qy  2222  -------VSE---------WSDESWSTATTASSYVTGPPYPKIRGKDSTQSAP          2258
Db  2331  RTVGLSESTISEALQQLAIKTFGQPPSSGDAGSSTGAGAABSGGPTSPGEPAPSETGSAS   2390

Qy  2259  AKRPTKKLGKSEFS-------------------------CSMSYTWTD             2282
Db  2391  SMPPLEGBGPGDPDLESDQVELQPPPQGGGVAPGSGSGSWSTCSEEDDTTVCCSMSYSWTG  2450

Qy  2283  -VISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQV   2341
Db  2451  ALITPCSPEEEKLPINPLSNSLLRYHNKVYCTTSKSASQRAKKVTFDRTQVLDAHYDSVL   2510

Qy  2342  RLAKEKASKVVGVMWDYDEVAAHTPSKSSAKSHITGLRGTDVR--SGAARKAVLDLQKCVE  2399
Db  2511  KDIKLAASKVSARLLTLEEACQLTPPHSARSKY-GFGAKEVRSLSGRAVNHIKSVWK--D   2567

Qy  2400  AGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKA   2459
Db  2568  LLEDPQTPIPTTIMAKNEVFCVDPAKGGKKPARLIVPDLGVRVCEKMALYDITQKLPQA   2627

Qy  2460  VMGDAYGF-VDPRTRVKRLLSMWS--PDAVGATCDTVCFDSITPEDIMVETDIYSAAKL   2516
Db  2628  VMGASYGFQYSPAQRVEYLLKAWAEKKDPMGFSYDTRCFDSTVTERDIRTEESIYQACSL   2687

Qy  2517  SDQHRAGIHTIARQLYAGGPMIAYDGRBIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQ   2576
Db  2688  PEEARTAIHSLTERLIVGGPMFNSKGQTCGYRRCRASGVLTTSMGNTITCYVKALACKA   2747

Qy  2577  AGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTS   2636
Db  2748  AGIVAPTMLVCGDDLVVISESQGTEEDERNLRAFTEAMTRYSAPPGGDPPRPEYDLELITS  2807

Qy  2637  CSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVL   2696
Db  2808  CSSNVSVALGPRGRRRYLTRDPTTPLARAAWETVRHSPINSWLGNIIQYAPTIWVRMVL    2867
```

```
Db   771  YFAI---------FFVAA-------WHIRGRVVPLTTYC---------LTGLWP         799
Qy   717  LVALLTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGPFAHLLPRCALVNSYLWQRW   776
Db   800  F--CLLLMALPRQAYAYDAPVHGQICVGLLLILTLFTLTPGYKTLLLGQC------LW--W  849

Qy   777  ------ENWFWNVTLRPER---FFLVLVCFPGATYDALVTFCVCHVALLCLTSS        821
Db   850  LCYLLTLGEAMIQEWVPPMQVRGGRDGIAWAVTIFCPGVVPDITKW------LLALLGP    902

Qy   822  A----ASFFGTDSRVRAHRMLVRLGKCHAWYSHYVLKFFLLVFGE-NGVFFYKHL--HGD   874
Db   903  AYLLRAALTHVPYFVRAHALIRVCALVKQLAGGRRYVQVALLALGRWTGTYIYDHLTPMSD  962

Qy   875  VLPNDFASKLPLQEP--FFPFEGKARVYRNEGRRLACGDTVDGLPVVARLGDLVFAGLAM   932
Db   963  WAASGLRDLAVAVEPIIFSPMEKKVIVWGAE--TRACGDILHGLPVSARLGQEILLG---  1017

Qy   933  PPD------GWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSYMGFV  987
Db  1018  PADGYTSKGWKLLAPITAYAQQTRGLLGAIVSMTCRDRTEQAGEVQILSTVSQSPLGTT   1077

Qy   988  CDNVLYTAHHGSKGRRLLAHPTGSIHPITVDAANDDQIYQPPCGAGSLTRCSCGETKGYLV  1047
Db  1078  ISGVLWTVTHHGAGNKTLAGLRGPVITQMYSSAEGDLVGWPSPPGTKSLEPCKCGAVDLYLV 1137

Qy  1048  TRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSCHVIGMFTAARNSGG---SV  1104
Db  1138  TRNADVIPARRRGDKRGALLSPRPISTLKGSSGGPVLCPRGHVVGLFRAAVCSRGVAKSI   1197

Qy  1105  SQIRVRPLVCAGYHPQYTAHATLDIKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQEKYE   1164
Db  1198  DFIPVETLDVVTRSPTFSDNS---TPPAVPQTYQVGYLHAPTGSGKSTKVPVAYAAQGYK   1254

Qy  1165  VLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYLT--GACSRNY   1222
Db  1255  VLVLNPSVAATLGFGAYLSKAHGINPNIRTGVRTVWTGEAITYSTYGKFLADGGCASGAY   1314

Qy  1223  DVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDE   1282
Db  1315  DIIICDECHAVDATSIIGIGTVLDQAETAGVRLTVLATATPPGSVVTPHPDIEEVGLGRE   1374

Qy  1283  GTIPFHGKKIKEENLKKGRHLIFEATKGHCDELANELARKGITAVSYYRGCDISKIP-EG  1341
Db  1375  GEIPFYGRAIPLSCIKGGRHLIFCHSKKKCDELAAALRGMGLNAVAYYRGLDVSIIPAQG   1434

Qy  1342  DCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKQRRGR   1401
Db  1435  DVVVVATDALMTGYTGDFDSVIDCNVAVTQAVDFSLDPTFTITTQTVPQDAVSRSQRRGR   1494

Qy  1402  TGRGRAGIYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAGTILDYRTQPGLPA   1461
Db  1495  TGRGRQGTYRYVSTGERASGMFDSVVLCECYDAGAAWYDLTPAETTVRLRAYFNTPGLPV   1554

Qy  1462  IGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLCHQVGYAAPNDAPRWQG   1520
Db  1555  CQDHLEFWEAVFTGLTHIDAHFLSQTKQAGENFAYLVAYQATVCAR---AKAPPSWDA   1610

Qy  1521  -----ARLGKK---PCGVLWRLDGADACPGEPSEV------TRY----QMCFTEVNTSG  1562
Db  1611  MWKCLARLKPTLAGPTPLLYRL------GPITNEVTLTHPGTKYIATCMQADLEVMTS-   1662

Qy  1563  TAALAVGVGVAMAYLAIDTFGGATCVRRCWSITSVPTGATVAPVDEEEI------VEECA  1616
Db  1663  TWVLAGGVLAAVAAYCLATGCVSIIGR----LHVNQRVVVAP--DKEVLYEBAFDEMEECA  1716

Qy  1617  SFIPL-EAMVAAIDKLKSTI------TTTSPFTLETALE----KLNTPLGPHAATILAII  1665
Db  1717  SRAALIEEGQRIAEMLKSKIQGLLQQASKQAQODIQPAMQASWPKVEQFWARHMWNFISGI  1776

Qy  1666  EYCCGLVTLPDNPFASCVFAFIAGITTLPLPHKIKMFLSLFGGAIASKLTDARGALAFMMA  1725
Db  1777  QYLAGLSTLPGNPAVASMMAFSAALTSPLSTSTTILLNIMGGWLASQIAPPAGATGFVVS   1836
```

```
CC        hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
CC        polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
CC        and NS5B proteins), NS5A mutant of the invention. Note: The present
CC        sequence is not shown in the specification but was created by the indexer
CC        using the HCV sequence appearing as ABG32451 and the information in claim
CC        1
XX
SQ   Sequence 3010 AA;

Query Match             21.2%;  Score 3263;  DB 5;  Length 3010;
Best Local Similarity   29.8%;  Pred. No. 2.4e-243;
Matches  970; Conservative 452; Mismatches 1183; Indels 648; Gaps 116;

Qy      8 TSVPVAPRTRKNKTQASYPVSIK-----------------TSVERGQRAKRKVQRDARPR 51
          |:||  : ||| |  |  |  : |                 ||| |   | | : |
Db      3 TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR 59

Qy     52 NYKIAGIHDGLQTLAQAALP-------------------------------AHGWGRQD 79
          | :|||||||||||||||||                             ::|||||||
Db     60 GRR-------QPIPKARQPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPTD 111

Qy     80 PRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG--- 136
          || :||||| || :||||||||||||||||||||||||||||||||||||||||
Db    112 PRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLP 170

Qy    137 WFGVHLFVVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC 195
          || :||:||| ::  ||:| |:|: |||  | | : | : | :  |     | :: | |
Db    171 GCSFSIFLLALLSCLTIPASAYEVRNV-SGVYHVTNDCSNASIVYEAADMIMHTPGCVPC 229

Qy    196 ADE--------CWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVG 250
          :| |        || |   || || ||:| |:| | ||||||||  :|:  : : | |
Db    230 VRENNSSRCWVALTPTLA-ARNASVPTTTIRRHVDLLVGAAALCSAMYVGDLCGSVFLVA 288

Qy    251 DWLV------RHWLIHIDLNETGTCYLEVPTGIDPGPL--GFIGW---MAGKVEAVIFLTKL 301
          ||||      :||:|  | ||:| |   | |||:|| |  || |      ::||| :|
Db    289 QLFTFSPRRHETVQ-DCN--------CSIYPGHVTGHRMAWDMMNWSPTAALVVSQL 337

Qy    302 ASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSIA 361
          :| || : : | : ||||| :|| || |:|| |    | | |:||::| |  | | | |
Db    338 L-RIPQAVVDMVAGAHWGVLAGLAYYSMVGNWAKVLIVMLLFAGVDGGTYV---TGGTMA 393

Qy    362 E-------------------------------FCSPLMIPCPCHSYLSE 379
          :                               : :|| | |  || ||
Db    394 KNTLGITSLFSPGSSQKIQLVNTNGSWHINRTALNCNDSLNTGFLAALFY---VHKFNSS 450

Qy    380 ------NVSEVICYSPKWTRPITLEYNNSISWYPY-----------TIPGARGC---- 416
               || | |  | | |  | | ||| | ||                |||| |
Db    451 GCPERMASCSPIDAFAQGW-GPITYNESHSSSDQRPYCWHYAPRPCGIVPAAQVCGPVVCF 509

Qy    417 ----------------MVKFKNNT------WGCCRIRNVPSY--------C 437
                          : ||:|         || | : ||
Db    510 TPSPVVVGTTDRFGVPTYSWGENETDVLLLNNTRPPQGNWFGCTWMNSTGFTKTCGGPPC 569

Qy    438 TMG--------------TDAVWNDTRNTYEACGVTPWLT-------TAWHNGSALKLAILQ- 477
          | |              || |||| ||| |  | |            ::| |:|  :| |
Db    570 NIGGIGNKTLTCPTDCFRKHPEATYTKCGSGPWLTPRCLVHYPYRLWHYPCTVNFTIFKV 629

Qy    478 --YPGSKE--MFKPHNWMSG-HLYFEGSDTPIVYFDPVNSTLLPPERWARLPGTPPVVR 532
            :|| |   |||  || |  :|    :|| ||  |   :: : |   | | || ||
Db    630 RMYVGGVEHRLEAACNWTRGERCNLEDRDR-------SELSPLLLSTTEWQVLPCS----- 678

Qy    533 GSWLQVPGQFYSDVKDLIATGLITKDKAWKNYQVLYSATGALSLTGVTKAVVLILLGLCG 592
          |:|   | | : |    | |  :||  |  | ||| ||| |  ::|  |   : : | ||
Db    679 ------------FTTLPALSTGLIHLHQNVDVQVLYGIGSAVVSFAIKWEYVLLLFLLAD 728

Qy    593 SKYLIAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYLR 652
          :|  :|  |   ||  | :| |        | | | ||: |  | | ::| |:  | |:
Db    729 ARVCACLWMMLL-----------------IAQAEAALENLVV------- 753

Qy    653 CRLRYAALLGFVPMAAGLPLTFVAAAAAQPDYDWWVR-LLVAGLVLWAGRNRGHRIALL 711
          |: :  |  ||   | |: |   |:     |  | |   |   |    | | | | |
Db    754 --LNAASVAG---AHGI-LSFLVFFCAA------WYIKGRLVPGAAY----------AL 790
```

RESULT 12

```
ABG32457
ID   ABG32457 standard; protein; 3010 AA.
XX
AC   ABG32457;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Hepatitis C virus Con 1 isolate polyprotein mutant #6.
XX
KW   HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutein;
KW   hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW   internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.
XX
OS   Hepatitis C virus.
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FT   Misc-difference 2173
FT                  /note="Wild-type Phe substituted by Ser"
XX
PN   WO200259321-A2.
XX
PD   01-AUG-2002.
XX
PF   16-JAN-2002; 2002WO-EP000526.
XX
PR   23-JAN-2001; 2001US-0263479P.
XX
PA   (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
PI   De Francesco R, Migliaccio G, Paonessa G;
XX
DR   WPI; 2002-599793/64.
XX
PT   New Hepatitis C virus (HCV) replicons comprising altered HCV
PT   NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
PT   internal ribosome entry site (IRES) region coding for one or more NS3,
PT   NS5A, or EMCV IRES mutations, respectively. The location of the mutations
PT   are detailed in the specification. Also included are (1) an expression
PT   vector comprising a nucleotide sequence coding for the altered nucleic
PT   acids, which is transcriptionally coupled to an exogenous promoter; (2) a
PT   recombinant cell human hepatoma cell comprising the altered nucleic acids
PT   ; (3) a recombinant cell produced by introducing into a human hepatoma
PT   cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
PT   replicon enhanced cell or which containing a functional HCV replicon; (5)
PT   an HCV replicon enhanced cells made in the method; and (6) measuring the
PT   ability of a compound to affect HCV activity. The HCV replicons and HCV
PT   replicon enhanced cells are useful in studying HCV replication and
PT   expression, and HCV and host cell interactions, producing HCV RNA and
PT   proteins, and providing a system for measuring the ability of a compound
PT   to modulate one or more HCV activities e.g. to discover drugs which may
PT   treat HCV mediated diseases such as liver failure, cirrhosis and
PS   Claim 1; Page: 69pp; English.
XX
```

```
Qy   2697 AVHFMEQMLFEDKLPETVTFDPWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRV 2756
          ||:|| :::: |  ||  | |: |  |:  ||    | :  | :||   |    : |  |
Db   2868 MTHFFSILMVQDTLDQNLNFEMYGSVYSVNPLDLPAIIERLHGLDAFSMHTYSHHELTRV 2927

Qy   2757 SQSLTDMWTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WHATSR---PLPDLDKT 2811
          |:|| |  ||||| | ||||||| ||||:||||||  || |  | |||    ||| ||
Db   2928 ASALRKLGAPPLRVWKSRARAVRASLISRGGKAAVCGRYLFNWAVKTKLKLITPLPEARLL 2987

Qy   2812 SVARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAI 2863
          :|||| ||  |   | || :|:|| :    |    | |:| | | :  |
Db   2988 DLSSWFT-----VGAGGDIFHSVSRARPRSL------LFGLLLFVGVGL 3027
```

```
Qy   712  VGPWPLVALLTLLHLVTPASAPDTEI---IGGLTIPPVVALVVMSRFGFPAHLLPRCALV   768
          |:|| || ||||  | |  ||   |       ||| |:|||||:|||:  : || || ||
Db   791  YGVWPL-LLLLALPPRAYAMDREMAASCCGGAVFVGLILILTLSPHYKLFLARL-------   842

Qy   769  NSYLWQRWENWFWNVTLRPERFFLVLVCFP-----GATYDALVTFCVCHVALL-CLTSSA   822
          ||:|  : |  |  | |  |||||||| |       || |||||| | | ||  || ||
Db   843  ---IW-----WLQYFITRAEAHLQVWI--PPLNVRGGRDAVILLTCAIHPELIFTITKIL   892

Qy   823  ASFFG----------TDSRVRAHR-----MLVRLGKCHAWYSHYVLKFFLLVFGENGVF   866
          |  ||             | |  |     ||||
Db   893  LAILGPLMVLQAGITKVPYFVRAHGLIRACMLVR----KVAGGHYVQMALMKLAALTGTY   948

Qy   867  FYKHL-------HGDVLPNDFASKLPLQEPFPFFEGKARVVRNEGRRLACGDTVDGLPVV   919
          |  |            |  ||:     | :   |        |   ||  | ||||||
Db   949  VYDHLTPLRDWAHAGL--RDLAVAV---EPVVFSDMETKVITWGADTAACGDIILGLPVS  1003

Qy   920  ARLGDLVFAGLA--MPPDGWAITAPPTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLG   977
          |||  ||| | :  ||  |: |  : :  |:   |  : ||||: | || |||| |||  |
Db  1004  ARRGREIHLGPADSLEGGQWRLLAPITAYSQQTRGLLGCIITSLTGRDRNQVEGEVQVVS  1063

Qy   978  SLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN-DQDI--YQPPCGAGSL  1034
          || | ||||  ||| | ::   | :| | | |: ::| :|    : ||      |   ||
Db  1064  TATQSFLATCVNGVCWTVYHGAGSKTLAGPKG---PITQMYTNVDQDLVGWGNQAPPGARSL  1120

Qy  1035  TRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMF  1094
          | |:|| | | ||  | ||:|  :|| |  | || | |   ||:|||||||:||||||:
Db  1121  TPCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSSGPLLCPSGHAVGIF  1180

Qy  1095  TAA---RNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILLAPTGSGKS  1151
          |||     | | | | ||:||| | || | |  |      ::   :: ||  ||:||||
Db  1181  RAAVCTRGVAKAVDFVPVESMETTMRSPVFTDNS---SPPAVPQTFQVAHLHAPTGSGKS  1237

Qy  1152  TKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYG  1211
          |||| |   ||||||:||| |   |  ||| ||  |:|||  |  |  | |  |  || ||
Db  1238  TKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGAPITYSTYG  1297

Qy  1212  MYLT-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPGVIPT   1269
          |   ||| | ||:|||||||:|   ||| |   |     ||  |  ||| ||||  ||
Db  1298  KFLADGGCSGGAYDIIICDECHSTDSTTILGIGTVLDQAETAGARLVVLATATPPGSVTV  1357

Qy  1270  PHANITEIQLTDEGTIPHGKKIKEENLKGRHLIFEATKKGHCDELANELARKGITAVSY   1329
          || || :| | |  |  |  |    |||||||||    :  | |:| || | :||| |
Db  1358  PHPNIEEVALSSTGEIPFYGKAIPIETIKGGRHLIFCHSKKKCDELAAKLSGLGLNAVAY  1417

Qy  1330  YRGCDISKIP-EGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVC  1388
          |||   | || |||  ||||||| |||:||||||:||: :  |  |||| ||   |
Db  1418  EDVVCCSMSYTWTGALIITPCAAEETKLPINALSNSLLRHHNLVYATTSRSASLRQKKVTF  1477

Qy  1389  GVSAIVKGQRRGRTGRGRAGIYYYVDGSCCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQT  1448
          | |: |||:|||||||||||||  ||   | ||| | | | | |   ||| |     |
Db  1478  YRGLDVSVIPTSGDVIVVATDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTV  1537

Qy  1449  ILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNVLLTAAQLQHCHQY   1507
          :| |   |||||   |:|: |   ||  ||| |  |   | | ||  ||  || | | ||
Db  1538  RLRAYLNTPGLPVCQDHLEFWESVFTGLTHIDAHFLSQTKQAGDNFPYLVAYQATVCARA  1597

Qy  1508  GYAAPNDAPRWQG-ARLGKK----PCGVLWRLDGADACPGEPSEVTRYQM-CFT-----EVN  1559
          || |  |   | :   |  |    |  ||| |||   | || | | ||      ||      ||
Db  1598  QAPPSWDQMWKCLIRLKPTLHGPTPLLYRL-GAVQNEVTTHPITKYIMACMSADLEVV    1656

Qy  1560  TSGTAALAVGVGAMAYLAIDTFGATCVRRCWSIITSVPTGATVAPVVDEEI-----VE  1613
          ||    | ||||      |  ||  |       |  | | | |    |          :|
Db  1657  TS-TWYLVGGVLAALAAYCLTTGSVVIVGR------IILSGKPAIIPDREVLYREFDEME  1709

Qy  1614  ECASFIP-LEAMVAAIDKLK-----STIT---TTSPFTLETALEKLNTFLGPHAATI  1661
          ||| |||     |||  ::|     |  |     ||| :|| | |:| | ||||  |
Db  1710  ECASHLPYIEQGMQLAEQFKQKAIGLLQTATKQAEAAAP-VVESKWRTLEAFWAKHMWNF  1768

Qy  1662  LAIIBYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALA  1721
          ::|   ||  || |  ||     |     |:  ||| : | ::  : ||| :  | ||:|
Db  1769  ISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTQHTLLFNILGGWVAAQLAPPSAAASA  1828

Qy  1722  FMMAGAAGTALGTWTISVGF---VEDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYS  1778
          |: ||  ||| |  ||| ||       | ||    ||  |  || |||| ||||||| :
Db  1829  FVGAGIAGAAVG---SIGLGKVLVDILAGYGAVGAGALVAFKVMSGEMPSTEDLVNLLPA  1885
```

```
Db  1779  AFNPAAGVGVLSACAMFALTTAGPD---HWPNRLLTMLARSNTVCNEYFIATRDIRRK  1834

Qy  1886  ILLSPGALVVGVV--CAALLRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAAR  1943

Db  1835  IIGILEASTPWSVISACIRWLHTPTEDDCGLI---AWGLEIWQYVCNFFVICFNVLKAGV  1891

Qy  1944  VTQILSSLT--ITQLLKRLHQWINEDCSTPCSGSWLRDVWDWICI----VLTDFKTWL  1995

Db  1892  QS--MVNIPGCPYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSN  1949

Qy  1996  QSKLLPRLPGVPFFSCQRGYKGVWRGDGIMQTTCPCGAQITGHVKNGSMRIV-GPRTCSN  2054

Db  1950  YWRGAVPVNARLCGSARPDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFT  2007

Qy  2055  TWHGTFPINAYTTGPCTPSPAPNYSRALWRVAAEEYVEVTRVGDFHYTGMTDNVKCPC  2114

Db  2008  QVP----------------------------------------PTLRAAVA-------VDGVQVCYLGEPKTPWTTS  2038

Qy  2115  QVPAPEFFTEVDGVRLHRYAPACKPLLREEVTFLVGLNQYLVGSQLPC---EPB-------  2165

Db  2039  ACCYGPDGKGKTVKL--PFRVDGHTPGVRMQLNLRDALETND-------CNSTNN  2084

Qy  2166  ------PDVAVLTFMLTDPSHITAETAKRLARGSPPSLASSSASQLSAPSLKATCTTRHD  2220

Db  2085  TPSDEAAVSALVFKQEL--------------RRTNQLL---EAISAGVDTTKLPAPSIEEVVRK  2132

Qy  2221  SPPDADLIEANLLWRQEMGGNITRVESENKVVILDSFEPLQAEBEDEREVSVPA--EILRRS  2278

Db  2133  RQFRARTGSLTLPPPPRSVPGVSCP-----------ESLQRSD-----------PLEGPSNLPPSP  2176

Qy  2279  RKF-----------------------PRAMPIWARPDYNPPLLESWKDPDYVPPVVHGCPLP-PAKAPPIP  2325

Db  2177  P-------------------------------VLQLAMPMPLLGAGCNPF------TAIGCAMTETGGGPDDLP  2214

Qy  2326  PRRRKRTVVLSESTVSSALAELA--TKTFGSSESSAVDSGTATASPDQPSDDGDAGSDVE  2383

Db  2215  SY---PP-----KKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLG  2268

Qy  2384  SYSSMPPLEGEGPDPDLSDGSWSTVSBEAS---------------------------:|  2413

Db  2269  KSBFSCSMSYTWTD-VISFKTASKVLSATRAITSGFLKQRSLYVTEPRDAELRKQKVTI  2327

Qy  2414  EDVVCCSMSYTWTGALIITPCAAEETKLPINALSNSLLRHHNLVYATTSRASLRQKKVTF  2473

Db  2328  NRQPLFPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGGA  2387

Qy  2474  DRLQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKF-GYGAKDVRN-LS  2531

Db  2388  RKAVLDLQK----CVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRC  2443

Qy  2532  SKAVNHIRSVWKDLLEDTETPI---DTTIMAXNEVFCVQPEKGGRKPARLIVFPDLGVRV  2588

Db  2444  VEKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMWSPD--AVGATCDTVCFDSTIT  2500

Qy  2589  CEKMALXDVVSTLPQAVMGSSYGFQYSPGQRVEFLVNAWKAKKCPMGPAYDTRCFDSTVT  2648

Db  2501  PEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSS  2560

Qy  2649  ENDIRVEBESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSC  2708

Db  2561  SNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASHMKVMGAP  2620

Qy  2709  GNTLTCYLKAAAACRAKLQDCTMLVCGDDLVICSSAGTQBEDEASLRAFTEAMTRYSAP  2768

Db  2621  QDCVPQPKYSLEELTSCSSNVTSGITKSGKPYFLTRDPRIPLGRCSAEGLGYNPSAAWI  2680

Qy  2769  PGDPPKPEYDLELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPVNSWL  2828

Db  2681  GYLIHHYPCLMVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGI  2740
```

```
Db    2829  GNIIMYAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGL 2888

Qy    2741  EAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WH 2799
            |||::|| :: :||||| ||| :||||| |: ||:| :::|| ::|| |:|| ||
Db    2889  SAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGRAATCGKYLFNWA 2948

Qy    2800  ATSR----PLP-----DLDKTSVARYTTFNYCDVYSPEPGDVFITPQRRLQKFLVKYLAVI 2850
            ::: :| |:| ||:| :: ::
Db    2949  VRTKLKLTPIPAASQLDLSSWFVAGYS--------GGDIYHSLSRARPRW--------F 2991

Qy    2851  VFALGLIAVGLAI 2863
            :| |||||
Db    2992  MWCLLLLSVGVGI 3004

RESULT 13
ABG32451
ID    ABG32451 standard; protein; 3010 AA.
XX
AC    ABG32451;
XX
DT    15-NOV-2002  (first entry)
XX
DE    Hepatitis C virus Con 1 isolate polyprotein.
XX
KW    HCV; Con 1; adaptive mutation; liver failure; cirrhosis;
KW    hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW    internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.
XX
OS    Hepatitis C virus.
XX
XX
PN    WO200259321-A2.
XX
PD    01-AUG-2002.
XX
PF    16-JAN-2002; 2002WO-EP000526.
XX
PR    23-JAN-2001; 2001US-0263479P.
XX
PA    (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
PI    De Francesco R, Migliaccio G, Paonessa G;
XX
DR    WPI; 2002-599793/64.
DR    N-PSDB; ABK91411.
XX
PT    New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT    NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT    ribosome entry site (IRES) region, useful in studying HCV replication and
PT    expression.
XX
PS    Claim 1; Page 34-36; 69pp; English.
XX
CC    The invention relates to nucleic acid molecules comprising altered HCV
CC    NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC    internal ribosome entry site (IRES) region coding for one or more NS3,
CC    NS5A, or EMCV IRES mutations, respectively. The location of the mutations
CC    are detailed in the specification. Also included are (1) an expression
CC    vector comprising a nucleotide sequence coding for the altered nucleic
CC    acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC    recombinant cell human hepatoma cell comprising the altered nucleic acids
CC    ; (3) a recombinant cell human hepatoma cell produced by introducing an
CC    HCV replicon enhanced cell or which containing a functional HCV replicon; (5)
CC    replicon enhanced cell or which containing a functional HCV replicon; (5)
CC    an HCV replicon enhanced cells made in the method; and (6) measuring the
CC    ability of a compound to affect HCV activity. The HCV replicons and HCV
CC    replicon enhanced cells are useful in studying HCV replication and
CC    expression, and HCV and host cell interactions, producing HCV RNA and
CC    proteins, and providing a system for measuring the ability of a compound
CC    to modulate one or more HCV activities e.g. to discover drugs which may
CC    treat HCV mediated diseases such as liver failure, cirrhosis and
CC    hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
CC    polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
```

```
CC    and NS5B proteins) used as a basis for the adaptive mutations of the
CC    invention
XX
SQ    Sequence 3010 AA;

Query Match             21.2%;  Score 3262;  DB 5;   Length 3010;
Best Local Similarity   29.8%;  Pred. No. 2.9e-243;
Matches  970;  Conservative  452;  Mismatches 1183;  Indels  648;  Gaps  116;

Qy      8  TSPVPAPRTRKNKQTQASYPVSIK-----------------------TSVERGQRAKRKVQRDARPR  51
           |:|:| ||:|| :|                                  ::||
Db      3  TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR  59

Qy     52  NYKIAGIHDGLQTLAQAALP------------------------------AHGWGRQD  79
                                                              :||||
Db     60  GRR-------QPIPKARQPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPTD  111

Qy     80  PRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVRPVCQIVRLLEDGVNWATG----  136
           ||:|||||| : |||| ::|| ||:|:||||| || || |:|||||||||
Db    112  PRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNVATGNLP  170

Qy    137  WFGVHLFVVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC  195
           |||| ||| : |||| ||:||| ||| :|| |::
Db    171  GCSFSIFLLALLSCLTIPASAEVRNV-SGVHVTNDCSNASIVYEAADMIMHTPGCVPC  229

Qy    196  ADE------CWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVG  250
                    |: :|: :|: :|| |::| | |:: :| ::::
Db    230  VRENNSSRCWVALTPTLA-ARNASVPTTIRRHVDLLVGAAALCSAMYVGDLCGSVFLVA  288

Qy    251  DWLV---RHWLIHIDLNETGTCYLEVPTGIDPGFL--GFIGW---MAGKVEAVIFLTKL  301
           |: | :| |::|| : |::: :| || :|| :
Db    289  QLFTFSPRRHETVQ-DCN-------CSIYPGHVTGHRMAWDMMMWSPTAALVVSQL  337

Qy    302  ASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSIA  361
           |: ||:| :| :|:| ::: ||| : ||: |: :|: :|| ::
Db    338  L-RIPQAVVDMVAGAHWGVLAGLAYYSMVGNWAKVLIVMLLFAGVDGGTYV---TGGTMA  393

Qy    362  E----------------------------------FCSPLMIPCPCHSYLSE  379
           :
Db    394  KNTLGITSLFSPGSSQKIQLVNTNGSWHINRTALNCNDSLNTGFLAALFY---VHKFNSS  450

Qy    380  -------NVSEVICYSPKWTRPITLEYNNSISWYPY----------------TIPGARGC-  416
                  :   :::| :::: :|  | ::                       :
Db    451  GCPERMASCSPIDAFAQGW-GPITYNESHSSDQRPYCWHYAPRPCGIVPAAQVCGPVYCF  509

Qy    417  -------------------------------------MVKFKNNT--------WGCCRIRNVPSY------C  437
                                                 |::| ||         ||
Db    510  TPSPVVVGTTDRFGVPTYSWGENETDVLLLNNTRPPQGNWFGCTWMNSTGFTKTCGGPPC  569

Qy    438  TMG---------TDAVNDTRNTYEACGVTPWLT-----------TAWHNGSALKLAILQ-  477
           |:|                                         ||
Db    570  NIGGIGNKTLTCPTDCFRKGHPEATYTKCCGSGPWLTPRCLVHYPYRLMHYPCTVNFTIFKV  629

Qy    478  --YPGSKE--MFKPHNWMSG-HLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVR  532
             |||:|   :||:| ||:|  |:::|||::|::::::::|::: | :
Db    630  RMVYGGVEHRLEAACNWTRGERCNLEDRDR------SELSPLLLSTTEWQVLPCS-----  678

Qy    533  GSWLQVPQGFYSDVKDLATGLITKDKAWKQNYQVLYSATGALSLTGVTTKAVVLLLGLCG  592
           |: | |  :: : || |::| :|:|:|: | | |
Db    679  ---------FTTLPALSTGLIHLHQNVVDVDQYLYGIGSAVVSFAIKWEYVLLLFLLLAD  728

Qy    593  SKYILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYLR  652
           |:  | :| ||: |  ||:|: | |: :| :| |: :
Db    729  ARVCACLWMMLL-----------------------------------IAQAEAALENLVV------  753

Qy    653  CRLRYAALLGFVPMAAGLPLTFFVAAAAAQPDYDWVR-LLVAGLVLWAGRNRGHRIALL  711
           |: ::| :| | |: :|       | | || :: :|
Db    754  --LNAASVAG----AHGI-LSFLVFFCA----WYIKGRLVPGAAY-----------AL  790

Qy    712  VGPWPLVALLTLLHLVTPASAFDTEI---IGGLTIPPVVALVVMSRFGFAHLLPRCALV  768
           ||| |:|| |:|  |:           |||:| |||: :|: :||  | ||:
Db    791  YGVWPL--LLLLALPPRYAMDREMAASCGGAVFVGLILLTLSPHYKLFLARL------  842
```

```
Db  1886 ILSPGALVVGVV--CAAILRRHVGPGEGAVQVWNRLLIAFASRGNHVSPTHYVPESDAAAR 1943
         :|  ||||:      |||   ||::      :: ::|  |
Qy  1835 IIGILEASTPWSVISACIRWLHTPTEDDCGLI----AWGLEIWQYVCNFFVICFNVLKAGV 1891

Db  1944 VTQILSSLT----ITQLLKRLHQWINEDCSTPCSGSWLRDVWDWICT----VLTDFKTWL 1995
         :|:||::|      ||:|||||:||:|||||||:|||||||||||||      ||:|||||
Qy  1892 QS--MVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGABLIFSVENGFAKLYKGPRTCSN 1949

Db  1996 QSKLLPRLPGVPFFSCQRGYKGVWRGDGIMQTTCPCGAQITGHVKNGSMRIV-GPRTCSN 2054
         |  :||: |  || ||:|||||||| | ::| ||| | ||   |:|| |:|  ||||||
Qy  1950 YWRGAVPVNARLCGSARPDPT-DWTSLVVNYGRVDYCYEKMGDHIFVTAVSSPNV-CPT 2007

Db  2055 TWHGTFPINAYTTGPCTPSPAPNYSRALWRVAAEEYVEVTRVGDFHYVTGMTDNVKCPC 2114
         ::  | ||||| | | |||||||||||||| |||||||||||||||||| || |:|| |
Qy  2008 QVP--------------------PTLRAAVA--------VDGVQVQCYLGPKTPWTTS 2038

Db  2115 QVPAPEFTEVDGVRLHRYAPACKPLLREEVTFLVGLNQYLVGSQLPC---EPE------- 2165
         ||||||||||||| :||| |||  ::|:||||:|||||||||:||      ||
Qy  2039 ACCYGPDGKGKTVKL--PFRVDGHTPGVRMQLNLRDALETND---------CNSTNN 2084

Db  2166 -----PDVAVLTSMLTDPSHITAETAKRLARGSPPSLASSSASQLSAPSLKATCTTRHD 2220
             :|:|| |||: || ||    |    :
Qy  2085 TPSDEARVSALVFKQEL-----RRTNQLL----EAISAGVDTTKLPAPSIEEVVVRK 2132

Db  2221 SPDADLIEANLLWRQEMGGNITRVESENKVVILDSFEPLQAEBDEREVSVPA--EILRRS 2278
         : |: |:: |  | :       ||::||::  |: |  |        :|
Qy  2133 RQFRARTGSLTLPPPPRSVPGVSCP----------ESLQRSD-------- 2176

Db  2279 RKP------PRAMPIWARPDYNPPLLESWKDPDYVPPVHGCPLP-PAKAPPIP 2325
         ||:
Qy  2177 P-------------------VLQLAMPPLLGAGECNPF----TAIGCAMTETGGGPDDLP 2214

Db  2326 PPRRKRTVVLSESTVSSALAELA--TKTFGSSESSAVDSGTATASPDQPSDDGDAGSDVE 2383
         :: ::: | |:|||:||:      ||| |||||:||
Qy  2215 SY---PP---KKEVSEWSDESWSTATTASSVYTGPPYPKIRGKDSTQSAPAKRPTKKKLG 2268

Db  2384 SYSSMPPLEGEPGDPDLSDGSWSTVSSEEAS----------------- 2413
         ||:| ||| |||  |:|| |||||:||||:|
Qy  2269 KSEFSCCSMSYTWTD-VISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTI 2327

Db  2414 EDVVCCSMSYWTGALITPCAAEETKLPINALSNSLLRHHNLVYATTSRSASLRQKKVTF 2473
         :|| |||||| ||::  | |||:|||||| | |||| | :||||:||||||| |:|||:
Qy  2328 NRQPLFPPSYHKQVRLAKEKASKVVGVWMDYDEVAAHTPSKSAKSHITGLRGTDVRSGAA 2387

Db  2474 DRLQVLLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKF-GYGAKDVRN-LS 2531
          |:|:||: |:  || ||||||||| || ||:|||| |||||||||   || ||||  :|
Qy  2388 RKAVLDLQK----CVRAGEIPSHYRQTVIVPKEEVVKTPKQPTKKPPRLISYPHLEMRC 2443

Db  2532 SKAVNHIRSVWKDLLEDTETPI---DTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRV 2588
         ||||:||||||||||||||||    ||||:||:||| ||||:| ||||||||||||||||
Qy  2444 VEKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMWSPD--AVGATCDTVCFDSTIT 2500

Db  2589 CEKMALYDVVSTLPQAVMGSSYGFQYSPGQRVEFLVNAWKAKKCPMGFAYDTRCFDSTVT 2648
         ||||| |||| ||||||||||||||||| ||||||| ||||||||||||||| |||||:|
Qy  2501 PEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTSS 2560

Db  2649 ENDIRVEBSIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSC 2708
         ::|| ||:|||  ||| ||||||||| ||||||||| | ||| |||||||:|||:|:|
Qy  2561 SNSLTCWLKVNAAABQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAP 2620

Db  2709 GNTLTCYLKAAAACRAAKLQDCTMLVCGDDLVVICESAGTQEDEASLRAFTEAMTRYSAP 2768
          |:||||||  ||| || |  | ||||||| ||| |||| :|||| |||| |||:||||
Qy  2621 QDCVPQPKYSLEELTSCCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWI 2680

Db  2769 PGDPPKPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWL 2828
          || ||||||| |:|||||||:| | |||||||||||| :|| || |||| ||||||:||
Qy  2681 GYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGI 2740

Db  2829 GNIIMYAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGL 2888
         | :: : | ||  :|  |:||||| | ||||  || ||| |  ||| ||| |:|:|:|
Qy  2741 EAFSVRYTNABILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WH 2799
```

```
Qy   769 NSYLWQRWENWFWNVTLRPERFFLVLVCFP------GATYDALVTFCVCHVALL-CLTSSA 822
         :: |: |||:      |||   |||| ||        ||||  | ||:||  :|
Db   843 ---IW-----WLQYFITRAEAHLQVWI--PPLNVRGRGRDAVILLTCAIHPELIFFITKIL 892

Qy   823 ASFFG-------TDSRVRAHR-----MLVRLGKCHAWYSHYLKFFLLVFGENGVF 866
         | |:|        |||| | |      |||| | :|  |:|||| |||||| ||
Db   893 LAILGPLMVLQAGITKVPYFVRAHGLIRACMLVR----KVAGGHYVQMALMKLLAALTGTY 948

Qy   867 FYKHL-------HGDVLPNDFASKLPLQEPFFPFEGKARVYRNEGRRLACGDTVDGLPVV 919
          :|:|        ||| | |::| |   |  | |  ::| | | |   |||| : |||||
Db   949 VYDHLTIPLRDWAHAGL--RDLLAVAV---EPVVFSDMETKVITWGADTAACCGDIILGLPVS 1003

Qy   920 ARLGDLVFAGLA--MPPDGWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLG 977
         |:  : : | |   :  || |:   : : | ||     || |:|||||||||||| |||
Db  1004 ARRGREIHLGPADSLEGQGWRLLAPITAYSQQTRGLLGCIITSLTCGRDRNQVEGEVQVVS 1063

Qy   978 SLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN-DQDI--YQPPCCAGSL 1034
         : |:|:  |||::| ||||||| |  |  |     |  | : |  |:     |||  ||||
Db  1064 TATQSFLATCVNGVCWTVYHGAGSKTLAGPKG---PITQMYTNVDQDLVGWQAPPGARSL 1120

Qy  1035 TRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMF 1094
         | ||||  | |||||   :  | : | :   | | ||  |:|||| ||||||||::|:|
Db  1121 TPCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIF 1180

Qy  1095 TAA---RNSGGSVSQIRVRPLVCAGYHPQTAHATLDTKPTVPNEYSVQILIAPTGSGKS 1151
          ||   | | | | :: :|  |  |::| ||| |  | | ||| |||:: |||||||||
Db  1181 RAAVCTRGVAKAVDFVPVESMETTMRSPVFTDNS---SPPAVPQTFQVAHLHAPTGSGKS 1237

Qy  1152 TKLPLSYMQEKYEVLVLNPSVATTASMPKIYMHATYGVNPNCYFNGKCTNTGASLTYSTYG 1211
         |:: : | :: |:|||||||||| :   |: : |:||::  |::| ||:|| |: ||||||
Db  1238 TKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGAPITYSTYG 1297

Qy  1212 MYLT-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPT 1269
         ::|  |:| |  ||:||||||| || || |||| ||:   : |:|:|||||||||| :|
Db  1298 KFLADGGCSGGAYDIIICDECHSTDSTTILGIGTVLDQAETAGARILVVLATATPPGSVTV 1357

Qy  1270 PHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKGHCDELANELARKGITAVSY 1329
         || ||  |   |||||| ||| |  |  |||||||| :  | |||||:|| : |: || |
Db  1358 PHPNIEVALSSTGEIPFYGKAIPIETIKGGRHLIFCHSKKKCDELAAKLSGLGLNAVAY 1417

Qy  1330 YRGCDISKIP-EGDCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVC 1388
         |||  |  || :|| :||||||  |:|||||||:|| | | : | ||||||||:  : |
Db  1418 YRGLDVSVIPTSGDVIVATDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTV 1477

Qy  1389 GVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQT 1448
           : ::|||||||||||| ||| :| |  |||||  : ||| |:|| ||  || |:|:
Db  1478 PQDAVSRSQRRGRTGRGRMGIYRFVTPGERPSGMFDSSVLCBCYDAGCAWELTPAETSV 1537

Qy  1449 ILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLQLCHQY 1507
          | :| |:|||| :    :| |:: | ||| ||||::|  :: || | :  || :|  :
Db  1538 RLRAYLNTPGLPVCQDHLEFWESVFTGLTHIDAHFLSQTKQAGDNFPYLVAYQATVCARA 1597

Qy  1508 GYAAPNDAPRWQG-ARLGKK-----PCCGVLWRLDGADACPGPEPSEVTRYQM-CFT---EVN 1559
                 |      |      ||||: | ||  | || | | :| ||
Db  1598 QAPPPSWDQMWKCLIRLKPTLHGPTPLLYRL-GAVQNEVTTTHPITKYIMACMSADLEVV 1656

Qy  1560 TSGTRALAVGGVAMAYLAIDTFGGATVCRRCWSITSVPTGATVAPVVDEEBI-----VE 1613
         || | |||| |::  |||:    : :: |   :|:  |   :|  :|:  |:    |
Db  1657 TS-TWLVGGVLAALLAAYCLTTGSVVIVGR--------ILLSGKPAIIPDREVLYREFDEME 1709

Qy  1614 ECASFIP-LEAMVAAIDKLK-------STIT----TTSPFTLETALEKLNTFLGPHAATI 1661
         ||||:|| :|  : :  :|        |:||     |||| |:|| |||| | ||  |
Db  1710 ECASHLPYIEQGMQLAEQFKQKAIGLLQTATKQAERAAP-VVESKWRTLRAFWAKHMWNF 1768

Qy  1662 LAIIEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALA 1721
         :: | | |||:||| || |: ::|| |:|:|| | :  ::| :|||: |:|| || ::|
Db  1769 ISGIQYLAGLLSTLPGNPAIASLMAFTASITSPLTTQHTLLFNILGGWVAAQLAPPSAASA 1828

Qy  1722 FMMAGAGTALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYS 1778
         |: ||| |:|   : :|    :  :| |||||: | | || :| ||  | | :|:|
Db  1829 FVGAGIAGAAVG---SIGLGKVLVDILAGYGAGVAGALVAFKVMSGEMPSTEDLVNLLPA 1885

Qy  1779 AFNPAAGVVGVLSACAMFALTTAGPD----HWPNRLLTMLARSNTVCNEYFIATRDIRRK 1834
```

```
Db   2889  SAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCCKYLFNWA  2948
Qy   2800  ATSR----PLP-----DLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFIVKYLAVI  2850
Db   2949  VRTKLKLTPIPAASQLDLSSWFVAGYS--------GGDIYHSLSRARPRW------F     2991
Qy   2851  VPALGLIAVGLAI  2863
Db   2992  MWCLLLLSVGVGI  3004
```

RESULT 14
ABG32460
ID   ABG32460 standard; protein; 3010 AA.
XX
AC   ABG32460;
XX
DT   15-NOV-2002 (first entry)
XX
DE   Hepatitis C virus Con 1 isolate polyprotein mutant #9.
XX
KW   HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutein;
KW   hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW   internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.
XX
OS   Hepatitis C virus.
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   Misc-difference 2199
FT   /note= "Wild-type Ala substituted by Thr"
XX
PN   WO200259321-A2.
XX
PD   01-AUG-2002.
XX
PF   16-JAN-2002; 2002WO-EP000526.
XX
PR   23-JAN-2001; 2001US-0263479P.
XX
PA   (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
PI   De Francesco R, Migliaccio G, Paonessa G;
XX
DR   WPI; 2002-599793/64.
XX
PT   New Hepatitis C virus (HCV) replicons comprising altered HCV
PT   NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
PT   internal ribosome entry site (IRES) region, respectively. The location of the mutations
PT   ribosome entry site (IRES) region, useful in studying HCV replication and
PT   expression.
XX
PS   Claim 1; Page: 69pp; English.
XX

CC   and NSSB proteins), NS5A mutant of the invention. Note: The present
CC   sequence is not shown in the specification but was created by the indexer
CC   using the HCV sequence appearing as ABG32451 and the information in claim
XX   1
SQ   Sequence 3010 AA;

Query Match   21.2%; Score 3262; DB 5; Length 3010;
Best Local Similarity   29.8%; Pred. No. 2.9e-243;
Matches 970; Conservative 452; Mismatches 1183; Indels 648; Gaps 116;

Qy   8 TSPVPAPRTRKNKQTQASYPVSIK-----------------TSVERGQRAKRKVQRDARPR   51
Db   3 TNPKPQRKTKRNTNRR--PQDVKPPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR   59

Qy   52 NYKIAGIHDGLQTLAQAALP----------------------AHGWGRQD   79
Db   60 GRR--------QPIPKARQPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPTD 111

Qy   80 PRHKSRNLGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG---   136
Db 112 PRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLP 170

Qy 137 WFGVHLFVCLLS-LACPCSGARVTDPDTNTTLTNCCQRNQVIYCSPSTCLHEPGCVIC   195
Db 171 GCSFSIFLLALLSCLTIPASAYEVRNV-SGVYHVTNDCSNASIVYEAADMIMHTPGCVPC 229

Qy 196 ADE------CWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVG 250
Db 230 VRENNSSRCWVALTPTLA-ARNASVPTTTIRRHVDLLVGAAALCSAMVVGDLCGSVFLVA 288

Qy 251 DWLV----RHWLIHIDLNETGTCYLEVPTGIDPFGL--GFIGW--MAGKVEAVIFLTKL   301
Db 289 QLFTFSPRRHETVQ-DCN-----CSIYPGHVTGHRMAWDMMNWSPTAALVVSQL   337

Qy 302 ASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSIA 361
Db 338 L-RIPQAVDMVAGAHWGVLAGLAYYSMVGNWAKVLIVMLLFAGVDGGTYV--TGGTMA   393

Qy 362 E------------------------------------FCSPLMIPCPCHSYLSE   379
Db 394 KNTLGITSLFSPSGSSQKIQLVNTNGSWHINRTALNCNDSLNTGFLAALFY---VHKFNSS 450

Qy 380 ----------NVSEVICYSPKWTRPITLEYNNSISWYPY----------TIPGARGC---- 416
Db 451 GCPERMASCSPIDAFAQGW-GPITYNESHSSDQRPYCWHYAPRPCGIVPAAQVCGPVYCF 509

Qy 417 --------------------------MVKFKNNT------WGCCIRNVPSY---------C 437
Db 510 TPSPVVVGTTDRFGVPTYSWGENETDVLLLNNTRPPQGNWFGCTWMNSTGFTKTCGGPPC 569

Qy 438 TMG---------------------------TDAVWNDTRNTYEACCGVTPWLT--   477
Db 570 NIGGIGNKTLTCPTDCFRKGHPEATYTKCGSGPWLTPRCLVHYPYRLWHYPCTVNFTIFKV 629

Qy 478 --YPGSKE--MFKPHNWMSG-HLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVR 532
Db 630 RMYVGGVEHRLEAACNWTRGERCNLEDRDR--------SEISPLLLSTTEWQVLPCS----- 678

Qy 533 GSWLQVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLLLGLCG 592
Db 679 -----------FTTLPALSTGLIHLHQNVVDVQYLYGIGSAVVSPAIKWEYVLLLFLLAD 728

Qy 593 SKYILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLR 652
Db 729 ARVCACLWMLL------------------------IAQAEAALENLVV-------   753

Qy 653 CRLRYAALLGFVPMAAGLPLTFVAAAAQPDYDWVR-LLVAGLVLWAGRNGHRIALL   711
Db 754 --LNAASVAG----AHGI-LSFLVFFCAA-----WYIKGRLVPGAAY------AL   790

Qy 712 VGPWPLVALLTLLHLVTPASAFDTEI---IGGLTIPPVVALVVMSRFGFFAHLLPRCALV 768
```

CC   Claim 1; Page: 69pp; English.
CC   The invention relates to nucleic acid molecules comprising altered HCV
CC   NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC   internal ribosome entry site (IRES) region coding for one or more NS3,
CC   NS5A, or EMCV IRES mutations, respectively. The location of the mutations
CC   are detailed in the specification. Also included are (1) an expression
CC   vector comprising a nucleotide sequence coding for the altered nucleic
CC   acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC   recombinant cell human hepatoma cell comprising the altered nucleic acids
CC   ; (3) a recombinant cell human hepatoma cell produced by introducing into a human hepatoma
CC   cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC   replicon enhanced cell or which containing a functional HCV replicon; (5)
CC   an HCV replicon enhanced cells made in the method; and (6) measuring the
CC   ability of a compound to affect HCV activity. The HCV replicons and HCV
CC   replicon enhanced cells are useful in studying HCV replication and
CC   expression, and HCV and host cell interactions, producing HCV RNA and
CC   proteins, and providing a system for measuring the ability of a compound
CC   to modulate one or more HCV activities e.g. to discover drugs which may
CC   treat HCV mediated diseases such as liver failure, cirrhosis and
CC   hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
CC   polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A

```
Qy  AFNPAAGVGVLSACAMFALTTAGPD----HWPNRLLTMLARSNTVCNEYFIATRDIRRK 1834
        ||||  :|| || :    :            ||||||| :  ||     |  :  :||
Db  ILSPGALVVGVV--CAAILRRHVGPGEGAVQWMNRLLIAFASRGNHVSPTHYVPESDAAAR 1943

Qy  ILGILEASTPWSVISACIRWLHTPTEDDCGLI----AWGLEIWQYVCNFFVICFNVLKAGV 1891
        :  |   ||| |   :|   |           |  |   :| | | :      |   |
Db  VTQILSSLT----ITQLLKRLHQWINEDCSTPCSGSWLRDVWDWICT----VLTDFKTWL 1995

Qy  QS--MVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSN 1949
        |:   :||| ||| | ||:|| |    :|  |||  | :|    :| |  | ||| |||
Db  QSKLLPRLPGVPFFSCCQRGYKGWRGDGIMQTTCPCGAQITGHVKNGSMRIV-GPRTCSN 2054

Qy  YWRGAVPVNARLCGSARPDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CPT 2007
        :|     |  |  |  |       :||    :  ||| |  | :|  :|   |:  | |
Db  TWHGTFPINAYTTGPCTPSPAPNYSRALMRVAAEEYVEVTRVGDFHYVTGMTTDNVKCPC 2114

Qy  QVP---------------------------------------PTLRAAVA---------VDGVQVCYLGEPKTPWTTS 2038
        |:|                                                              |
Db  QVPAPEFFTEVDGVRLRHYAPACKPLLREEVTPLVGLNQYLVGSQLPC-------EPB---- 2165

Qy  ACCYGPDGKGKTVKL--PFRVDGHTPGVRMQLNLRDALETND--------------CNSTNN 2084
        ::   :                 |      :   ::  ::
Db  -----PDVAVLTSMLTDPSHITAETAKRLARGSPPSLTSSSASQLSAPSLKATCTTRHD 2220

Qy  TPSDEAAVSALVFKQEL-------RRTNQLL------EAISAGVDTTKLPAPSIEEVVVRK 2132
        : :  :|  :  | ||        || ||:
Db  SPDADLIEANLLWRQEMCGGNITRVESENKVVILDSFEPLQAEEDEREVSVPA--EILRRS 2278

Qy  RQFRARTGSLTLPPPPRSVPGVSCP-----------ESLQRSD--------PLEGPSNLPPSP 2176
        : :  |  |   :                      : |  |
Db  RKF---------------------PRAMPIWARPDYNPPLLESWKDPDYVPPVVHGCPLP-PAKAPPIP 2325

Qy  P----------------------------------VLQLAMPMPLLGAGECNPF------TAIGCAMETCGGPDDLP 2214
        |
Db  PPRRKRTVVLSESTVSSALAELA--TKTFGSSESSAVDSGTATASPDQPSDDGDAGSDVE 2383

Qy  SY-----PP----KKEVSEWSDESWSTATTASSYVTGPPYYPKIRGKDSTQSAPAKRPTKKKLG 2268
        |
Db  SYSSMPPLEGEFGDPDLSDGSWSTVSEEAS------------------------------- 2413

Qy  KSEFSCSMSYTWTD-VISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTI 2327
        | : : :|| || :  |  |:  ::|   ||    |  ::|    :: |: | |: :|
Db  EDVVCCSMSYTWTWGALITPCAEETKLPINALSNSLLRHHNLVYATTSRSASLRQKKVTF 2473

Qy  NRQPLFPPSYHKQVRLAKEKASKVVGVWMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAA 2387
        : :|:   |      ||   |::|  :| |  |  |  || |   | | : |
Db  DRLQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKF-GYGAKDVRN-LS 2531

Qy  RKAVLDLQK-----CVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRC 2443
        : | ::|       |   |    :||       ::| | ||   :  | | : |
Db  SKAVNHIRSVWKDLLEDTETPI--DTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRV 2588

Qy  VEKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMWSPD--AVGATCDTVCFDSTIT 2500
        :  :   | |   |     | |   ||  |  | :||: ||
Db  CEKMALYDVVSTLPQAVWGSSYFQYSPGQRVEFLVNAWKAKKCCPMGFAYDTRCFDSTVT 2648

Qy  PEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSS 2560
        :|  : | |:     |   | |  |  |:|  || | | |  |  |  | : | : |:
Db  ENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKQGNCGYRRCRASGVLTTSC 2708

Qy  SNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAP 2620
        :|  |   ||  |   || |   | ||||| : |     |::| || : | |:| | |
Db  GNTLTCYLKAAAACRAAKLQDCTMLVCGDDLVVICESAGTQEDEASLRAFTEAMTRYSAP 2768

Qy  QDCVPQPKYSLEBLTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWI 2680
        :||   | ||  | :|| ||| :| :| |  | :|  |  |  |  | :| || |:
Db  PGDPPKPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWL 2828

Qy  GYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGI 2740
        | :  |  |   |   |   :  :|    ||     |  |   :||: || : | || |
Db  GNIIMYAPTLWARMILMTHFFSILLAQELEKALDCQIYGACYSIEPLDLPQIIQRLHGL 2888
```

```
CC   hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
CC   polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
CC   and NS5B proteins), NS3 mutant of the invention. Note: The present
CC   sequence is not shown in the specification but was created by the indexer
CC   using the HCV sequence appearing as ABG32451 and the information in claim
CC   1
XX
SQ   Sequence 3010 AA;

Query Match             21.2%;  Score 3262;  DB 5;  Length 3010;
Best Local Similarity   29.8%;  Pred. No. 2.9e-243;
Matches  968;  Conservative 450;  Mismatches 1190;  Indels 642;  Gaps 114;

Qy      8 TSPVPAPRTRKNKQTQASYPVSIK------------------TSVERGQRAKRKVQRDARPR 51
          |:|| ||::|:|             :|:|
Db      3 TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR 59

Qy     52 NYKIAGIHDGLQTLAQAALP---------------------------------AHGWGRQD 79
          ||:|                                              ::
Db     60 GRR-------QPIPKARQPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPTD 111

Qy     80 PRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG--- 136
          || |||||   :::|           :|||||||
Db    112 PRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLP 170

Qy    137 WFGVHLFVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC 195
          |||||                    :||||||
Db    171 GCSFSIFLLALLSCLTIPASAYEVRNV-SGVYHVTNDCSNASIVYEAADMIMHTPGCVPC 229

Qy    196 ADE------CWVPANPYISHPSNWTGTDSFPLADHIDFVMGALVTCDALDIGELGCACVLVG 250
          |||                 |:|:
Db    230 VRENNSSRCWVALTPTLA-ARNASVPTTIRRHVDLLVGAAALCSAMYVGDLCGSVFLVA 288

Qy    251 DWLV---RHWLIHIDLNETGTCYLEVPTGIDPGFL--GFIGW--MAGKVEAVIFLTKL 301
          |||        ||
Db    289 QLFTFSPRRHETVQ-DCN----CSIYPGHVTGHRMAWDMMWWSPTAALVVSQL 337

Qy    302 ASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSIA 361
          : :               ||||
Db    338 L-RIPQAVVDMVAGAHWGVLAGLAYYSMVGNWAKVLIVMLLFAGVDGGTYV---TGGTMA 393

Qy    362 E----------------------------------------FCSPLMIPCPCHSYLSE 379
          :
Db    394 KNTLGITSLFSPGSSQKIQLVNTNGSWHINRTALNCNDSLNTGFLAALFY---VHKFNSS 450

Qy    380 -------NVSEVICYSPKWTRPITLEYNNSISWPY----------------TIPGARGC- 416
Db    451 GCPERMASCSPIDAFAQGW-GPITYNESHSSDQRPYCWHYAPRPCGIVPAAQVCGPVYCF 509

Qy    417 -----------------------MVKFKNNT-------WGCCRIRNVPSY-------C 437
          : :
Db    510 TPSPVVVGTTDRFGVPTYSWGENETDVLLLNNTRPPQGNWFGCTWMNSTGFTKTCGGPPC 569

Qy    438 TMG---------TDAVWNDTRNTYEACGVTPWLT--------TAWHNGSALKLAILQ- 477
          :
Db    570 NIGGIGNKTLTCPTDCFRKHPEATYTKCGSGPWLTPRCLVHYPYRLWHYPCTVNFTIFKV 629

Qy    478 --YPGSKE--MFKPHNWMSG-HLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVR 532
          :                                             :  :
Db    630 RMYVGGVEHRLEAACNWTRGERCNLEDRDR------SELSPLLLSTTEWQVLPCS---- 678

Qy    533 GSWLQVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVILLGLCG 592
          :
Db    679 ---------FTTLPALSLGLIHLHQNVDVQVLYGIGSAVVSFAIKWEYVLLLFLLLAD 728

Qy    593 SKYILAYLCYLSLCFGRASGYPLRPVLPSQSYILQAGWDVLSKAQVAPFALIFFICCYLR 652
          : |:
Db    729 ARVCACLWMMLL-------------------------IAQAEAALENLVV-------- 753

Qy    653 CRLRYAALLGFVPMAAGLPLTFVAAAAQPDYDWVR-LLIVAGLVLWAGRNRGHRIALL 711
          |:|:   |:
Db    754 --LNAASVAG----AHGI-LSFLVFFCAA------WYIKGRLVPGAAY--------AL 790
```

```
Qy   2741 EAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WH 2799
          |||:|  ||: |||:   : :||:|  ||: :|||  || :||: |
Db   2889 SAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAAATCGKYLRNWA 2948
Qy   2800 ATSR----PLP----DLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFPLVKYLAVI 2850
          :|           |  |:|
Db   2949 VRTKLKLTPIPAASQLDLSSWFVAGYS---------GGDIYHSLSRARPRW------F 2991
Qy   2851 VFALGLIAVGLAI 2863
          :: | |:|:||:|
Db   2992 MWCLLLLSVGVGI 3004
```

```
RESULT 15
ABG32453
ID   ABG32453 standard; protein; 3010 AA.
XX
AC   ABG32453;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Hepatitis C virus Con 1 isolate polyprotein mutant #2.
XX
KW   HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutein;
KW   hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW   internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.
XX
OS   Hepatitis C virus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1202
FT                   /note= "Wild-type Glu substituted by Gly"
XX
PN   WO200259321-A2.
XX
PD   01-AUG-2002.
XX
PF   16-JAN-2002; 2002WO-EP000526.
XX
PR   23-JAN-2001; 2001US-0263479P.
XX
PA   (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
PI   De Francesco R, Migliaccio G, Paonessa G;
XX
DR   WPI; 2002-599793/64.
XX
PT   New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT   NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT   ribosome entry site (IRES) region, useful in studying HCV replication and
PT   expression.
XX
PS   Claim 1; Page; 69pp; English.
XX
CC   The invention relates to nucleic acid molecules comprising altered HCV
CC   NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC   internal ribosome entry site (IRES) region coding for one or more NS3,
CC   NS5A, or EMCV IRES mutations, respectively. The location of the mutations
CC   are detailed in the specification. Also included are (1) an expression
CC   vector comprising a nucleotide sequence coding for the altered nucleic
CC   acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC   recombinant cell human hepatoma cell comprising the altered nucleic acids
CC   ; (3) a recombinant cell human hepatoma cell produced by introducing into a human hepatoma
CC   cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC   replicon enhanced cell or which containing a functional HCV replicon; (5)
CC   an HCV replicon enhanced cells made in the method; and (6) measuring the
CC   ability of a compound to affect HCV activity. The HCV replicons and HCV
CC   replicon enhanced cells are useful in studying HCV replication and
CC   expression, and HCV and host cell interactions, producing HCV RNA and
CC   proteins, and providing a system for measuring the ability of a compound
CC   to modulate one or more HCV activities e.g. to discover drugs which may
CC   treat HCV mediated diseases such as liver failure, cirrhosis and
```

```
Qy   712 VGPWPLVALITLLHLVTPASAFDTEI--IGGLTIPPVALVVMSRFGFFAHLLPRCALV------ 768
            |||||| ::    |    ||     ::   |  ||||  : ::  |    ||  :|||
Db   791 YGVWPL-LLLLLALPPRAYAMDREMAASCCGGAVFVGLILLTLSPHYKLFLARL-          842

Qy   769 NSYLWQRWENWFWNVTLRPERRFFLVLVCFP-----GATYDALVTFCVCHVALL-CLTSSA    822
            :    |||  || :         ::|  ::|     |  : :  | ||  | |
Db   843 ---IW--WLQYFITRABAHLQWNI--PPLNVRGGRDAVILLTCAIHPELIFTITKIL       892

Qy   823 ASFFG------TDSRVRAHR-----MLVRLGKCHAWYSHYVLKFFLLVFGENGVF         866
            |:||        | |  |        |||||
Db   893 LAILGPIMVLQAGITKVPYFVRAHGLIRACMLVR----KVAGGHYVQMALMKLAALTGTY    948

Qy   867 FYKHL-------HGDVLPNDFASKLPLQEPFPPFEGKARVVRNEGRRLACGDTVDGLPVV     919
            |:|:|        :|:| |  :: :  | :  |  |  ||  :
Db   949 VYDHLTPLRDWAHAGL--RDLAVAV---EPVVFSDMETKVITWGADTAACGDIILGLPVS    1003

Qy   920 ARLGDIVFAGLA--MPPDGWAITAPPTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLG    977
            ||  :||:: :   : |  | ::    :| |    :||:||  :||  |:| |  :|
Db  1004 ARRGREIHLGPADSLEGQGWRLLAPITAYSQQTRGLLGCIITSLTGRDRNQVEGEVQVVS   1063

Qy   978 SLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN-DQDI--YQPPCGAGSL   1034
            ||:||||:  |||: ||::|:|    ||   ||||| ||:|   |||    | | |
Db  1064 TATQSFLATCVNGVCWTVYHGAGSKTLAGPKG---PITQMYTNVDQDLVGWQAPPGARSL   1120

Qy  1035 TRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMF   1094
            | ||||  :||:|| ||||:  |:  :|| |  ||| ::   ||||||:|:|:|||| |
Db  1121 TPCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIF   1180

Qy  1095 TAARNSGGSVSQIRRVRPLVCAGYHPQVTAHATLDTKPTVPNEYSVQILIAPTGSGKSTKL   1154
            ||| ||||||:|||| |||:|||||| || || |||||| :|:|| :|:||||||||||
Db  1181 RAAVCTRGVAKAVDFVPVESMGTTMRSPVFTDNSSPPAVPQTFQVAHLHAPTGSGKSTKV   1240

Qy  1155 PLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYL   1214
            ||:|| :||:||:|||||| ||:     |     |   ::::         ||||||:
Db  1241 PAAYAAQGYKVLVLNPSVAATLGFAYMSKAHGIDPNIRTGVRTITTGAPITYSTYGKFL   1300

Qy  1215 T-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHA   1272
            :  |  |:  ||||||||: |||||||:|| |      || || | |||||||| |
Db  1301 ADGGCSGGAYDIIICDECHSTDSTTILGIGTVLDQAETAGARLVVLATATPPGSVTVPHP   1360

Qy  1273 NITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELARKGITAVSYRG   1332
            ||  :||  : :|| | || :   :|  ||||| ::    || |   |  |
Db  1361 NIEEVALSSTGEIPFYGKAIPIETIKGGRHLIFCHSKKKCDELAAKLSGLGNAVAYYRG   1420

Qy  1333 CDISKIP--EGDCVVVATDALCGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVS   1391
              :|:||   |||:||||||| |    |||  | |  : | | |   | ||  | | |
Db  1421 LDVSVIPTSGDVIVATDALMTGFTGDFDSVIDCNTCVTQVDFSLDPTFTIETTVPQD    1480

Qy  1392 AIVKGQRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTILD   1451
            | :  |||||||||| |:| | ||| : | |||  |:: || ||  ||  |  :
Db  1481 AVSRSQRRGRTGRGRMGIYRFVTPGERRPSGMFDSSVLCECYDAGCAWYELTPAETSVRLR   1540

Qy  1452 TYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLCHQYGYA    1510
             | ||||| :   |:   | |     || |      ||      :  |  :|  || |
Db  1541 AYLNTPGLPVCQDHLEFWESVFTGLTHIDAHFLSQTKQAGDNFPYLVAYQATVCARAQAP   1600

Qy  1511 APNDAPRWQG-ARLGKK---PCGVLWRLDGADACPGPEPSEVTRYQM-CFT--EVNTSG   1562
                  | | |     :     || | |||       | |       :       | |||
Db  1601 PPSWDQMWKCLIRLKPTLHGFPTLLYRL-GAVQNEVTTTHPITKYIMACMSADLEVVTS-   1658

Qy  1563 TAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI------VEECA   1616
            :| :  || | ||    ::  |  |         |        :            ||
Db  1659 TWVLVGGVLAALLAAYCLTTGSVVIVGR-------IILSGKPAIIPDREVLYREFDEMEECA   1712

Qy  1617 SFIP-LEAMVAAIDKLK------STIT----TTSPFTLETALEKLNTFLGPHAATLAI   1664
            :     :||     |         : |     :: |        ||  |  |  ::
Db  1713 SHLPYIEQGMQLAEQFKQKAIGLLQTATKQABAAAP-VVESKWRTLEAFWAKHMWNFISG   1771

Qy  1665 IEYCCGLVTLPDNPFASCVFAPIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMK   1724
              ||  : |  |:|  :|  |   |  |    | ::|       |     |  | | |
Db  1772 IQYLAGLSTLPGNPAIASLMAFTASITTSPLTTQHTLLFNILLGGWVAAQLAPPSAASAFVG   1831

Qy  1725 AGAAGTALGTWTSVGF---VFDMLGGYAAASTACLLTFKCLMGEWPTMDQLAGLVYSAFN   1781
            |  | |||| |||:||      :|||| |     |  |   |||  |  |  ||:| |
Db          IHHYPCLWVSRVLAVHFMEQML.FEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAF   2743
```

```
Qy  1832 AGIAGAAVG---SIGLGKVLVDILAGYGAGVGAGALVAFKVMSGEMPSTEDLVNLLPAILS  1888
            |  ||| ||    ||||:  | ::|| | ||| | |  | :|  ||  |||||||:| :
Db          PAAGVGVGLSACAMFALTTAGPD----HWPNRLLTMLARSNTVCNEYFIATRDIRRKILG   1837

Qy  1889 PGALVVGVV--CAAILRRHVGPGEGAVQWMNRLIAPASRGNHVSPTHVVPESDAAARVTQ   1946
            ||:|:|||    || |||||| | |  |  ||:|  |:| | |:||| ||:|  |  | |
Db  1838 ILEASTPWSVISACIRWLHTPTEDDCGLI---AWGLEIWQYVCNFFVICFNVLKAGVQS-  1893

Qy  1947 ILSSLIT---ITQLLKRLHQWINEDCSTPCSGSWLRDVWDWICT----VLTDFKTWLQSK  1998
            |  ::|    ||  |  | :  |  |   |  | |||  |||| |      | ||   |:|
Db  1894 -MVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWR  1952

Qy  1999 LLPRLPGVPFFSCQRGYKGVWRGDGIMQTTCPCGAQITGHVKNGSMRIV-GPRTCSNTWH  2057
            | |: ||||:|:| :|||| |||| || | | ||   | | :||    :  |||||| |
Db  1953 GAVPVNARLCGSARPDT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFTQVP  2010

Qy  2058 GTFPINAYTTGPCTPSPAPNYSRALWRVAAEEYVEVTRVGDFHYVTGMTTDNVKCPCQVP  2117
            |  | | |:|| | | ||  | | |  |  :|   |:|  |  :   |   |::   | |
Db  2011 ------------------PTLRAAVA-------VDGVQCYLGEPKTPWTTSACC         2041

Qy  2118 APEFFTEVDGVRLHRYAPACKPLLREEVTFLVGLNQYLVGSQLPC------EPE------  2165
            || :  :||| || |   |  || :|  | |||||   ||  |       |
Db  2042 YGPDGKGKTVKL--PFRVDGHTPGVRMQLNLRDALETND-----------CNSTNNTPS   2087

Qy  2166 --PDVAVLTSMLTDPSHITAETAKRRLARGSPPSLASSSASQLSAPSLKATCTTRHDSPD  2223
              ||  : ||: || :| : | :  |  |||||: | |  |: |  | | |  : ||
Db  2088 DEAAVSALVFKQEL------EAISAGVDTTKLPAPSIEEVVRKRQF                2135

Qy  2224 ADLIEANLLWRQEMGGNITRVESENKVVILDSFEPLQAEEDEREVSVPA--EILRRSRKF  2281
            ||:     ||:|::    | ||  :: :|  |  | | :       |      | |  |
Db  2136 RARTGSLTLPPPPRSVPGVSCP--------------ESLQRSD----------PLEGPSNLPPSPP--  2177

Qy          PRAMPIWARPDYNPPLLESWKDPDYVPPVVHGCPLP-PAKAPPIPPPR             2328
            |  |  |       |  :|| |    | | | |      | ||  | |
Db  2178 --------------VLQLAMEMPLLGAGECNPF-------TATGCAMTETGGGPDDLPSY-  2216

Qy  2329 RKRTVVLSESTVSSALAELA--TKTFGSSESSAVDSGTATASPDQPSDDGDAGSDVESYS  2386
            | | |||||| ||:| |||   || :|:||  |   |    |  | |  | |  |:||:|
Db  2217 --PP----KKEVSEWSDESWSTATTASSVVTGPPYPKIRGKDSTQSAPAKRPTKKLGKSE  2271

Qy  2387 SMPPLEGEPGDPDLSDGSWSTVSEEAS---------------EDV                  2416
            |:||||| | |  | || |||  ||::                :|
Db  2272 FSCSMSYTWTD-VISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQ  2330

Qy  2417 VCCSMSYTWTGALITPCAAEETKLPINALSNSLLRHHNLVYATTSRSASLRQKKVTFDRL  2476
            | | | ||| ||: | | ||| |  || :| |:|: |   :| ||: |  |:| ||| ||
Db  2331 PLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKA  2390

Qy  2477 QVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKF-GVGAKDVRN-LSSKA  2534
            ||||| ||||| |:|||||| ||  |||  ||| | | |||:|   |||| |  |  ||
Db  2391 VLDLQK-----CVEAGEIPSHYRQTVIVPKEEVVKTPQKPTKKPPRLISYPHLEMRCVEK  2446

Qy  2535 VNHIRSVWKDLLEDTETPI---DTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEK  2591
            |||:| ||| || |  |     ||:| |||||| |   | ||||| |  |  |  | | |
Db  2447 MYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMWSPD--AVGATCDTVCFDSTITPED  2503

Qy  2592 MALYDVVSTLPQAVMGSSYGFQXSPGQRVEFLVNAWKAKKCPMGFAYDTRCFDSTVTEND  2651
            | |||||| |||| ||||||| : |||||||| |||||| | |  | |||| |||:| ||
Db  2504 IMVETDIYSAAKLSDQHRAGIHTIARQLYAGGMIAYDGREIGYRCRSSGVYTTSSSNS    2563

Qy  2652 IRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNT  2711
            |  | ||  ||  | |||  |  ||  | ||||:| ||    |  | ||| | |||   |
Db  2564 LTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDC  2623

Qy  2712 LTCYLKAAAACRAAKLQDCTMLVCGDDLVICESAGTQEDEASLRAFTEAMTRYSAPPGD  2771
            |||:|||| |  | | |  | |||||  ||| ||| ||||| ||||  || |||||||
Db  2624 VPQPKYSLEELTSCSSNVTGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYL  2683

Qy  2772 PPKPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNI  2831
            || | || |::|||||| |:| ||||| || ||||| ||||| ||| | | ||  |||
Db  2684 IHHYPCLWVSRVLAVHFMEQML.FEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAF  2743
```

```
Db    2832 IMYAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGLSAF 2891

Qy    2744 SVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRGGAHAKLARFLL-WHATS 2802
                |:    ||     |:  :: ||||  ||   :||||     | :|  |     :  |
Db    2892 SLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVRT 2951

Qy    2803 R----PLP------DLDKTSVARYTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFA 2853
               :   |:|      || |:     ||:::  |    ::    |  |       ::
Db    2952 KLKLTPIPAASQLDLSSWFVAGYS----------GGDIYHSLSRARPRW-------FMWC 2994

Qy    2854 LGLIAVGLAI 2863
                | |:::||:|
Db    2995 LLLLSVGVGI 3004
```

Search completed: October 27, 2005, 15:41:35
Job time : 318 secs

```
          GenCore version 5.1.6
       Copyright (c) 1993 - 2005  Compugen Ltd.

OM protein - protein search, using sw model

Run on:    October 27, 2005, 15:33:26 ; Search time 73 Seconds
                                         (without alignments)
                                         3774.860 Million cell updates/sec

Title:             US-10-009-002-2
Perfect score:     15376
Sequence:          1 MPVISTQTSPVPAPRTRKNK........KYLAVIVFALGLIAVGLAIS 2864
Scoring table:     BLOSUM62
                   Gapop 10.0 , Gapext 0.5

Searched:          283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:       283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     PIR_79:*
               1:  pir1:*
               2:  pir2:*
               3:  pir3:*
               4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3273.5	21.5	3033	1	GNWVJ8	genome polyprotein
2	3262	21.2	3033	1	JQ1303	genome polyprotein
3	3237	21.1	3010	1	GNWVTW	genome polyprotein
4	3232	21.0	3010	1	A45573	genome polyprotein
5	3224.5	21.0	3011	1	S40770	genome polyprotein
6	3223	20.9	3010	1	GNWVCJ	genome polyprotein
7	3218	20.9	3010	1	GNWVTC	genome polyprotein
8	3204.5	20.8	3011	1	GNWVC3	genome polyprotein
9	3161	20.6	3010	1	S18030	genome polyprotein
10	3126.5	20.3	3011	1	GNWVCH	genome polyprotein
11	3119	20.3	3014	1	JC5520	genome polyprotein
12	2055	13.4	3005	2	T08841	polyprotein - dour
13	2029.5	13.2	2970	2	T08839	polyprotein - marm
14	1028.5	6.7	1435	2	T01075	polyprotein - hepa
15	871	5.7	386	2	S68016	ATPase/RNA helicas
16	811.5	5.3	876	2	PC2219	polypeptide - hepa
17	698.5	4.5	874	2	JQ0883	genome polyprotein
18	688	4.5	874	2	JQ0881	genome polyprotein
19	649.5	4.2	365	2	JQ0879	NS5 protein - hepa
20	645.5	4.2	216	2	S21137	genome polyprotein
21	634.5	4.1	365	2	JQ0880	genome polyprotein
22	530	3.4	492	2	PS0326	polyprotein - hepa
23	416	2.7	782	2	S18032	genome polyprotein
24	415	2.7	3988	1	GNWVBV	genome polyprotein
25	412	2.7	3898	1	S19875	genome polyprotein
26	406.5	2.6	3898	1	A44217	genome polyprotein
27	406.5	2.6	3898	2	S58295	polyprotein - hog
28	406	2.6	782	2	S19876	genome polyprotein
29	406	2.6	782	2	S18031	genome polyprotein
30	398	2.6	787	2	PN0677	hypothetical prote
31	396.5	2.6	640	2	JQ1584	genome polyprotein
32	396.5	2.6	3898	1	GNWVHB	genome polyprotein
33	396	2.6	3898	1	S57437	genome polyprotein
34	394.5	2.6	189	2	S60587	non-structural pro
35	377	2.5	3898	1	GNWVHC	genome polyprotein
36	361.5	2.4	411	2	PC2061	genome polyprotein
37	356.5	2.3	492	2	S41288	genome polyprotein
38	352	2.3	513	2	A44150	structural protei
39	350	2.3	513	2	PC1284	genome polyprotein
40	348	2.3	441	2	S12707	genome polyprotein
41	345	2.2	411	2	PC2060	genome polyprotein
42	342	2.2	550	2	JH0711	polyprocein - hepa
43	337	2.2	520	2	JQ1925	genome polyprotein
44	336.5	2.2	3411	1	GNWVY	genome polyprotein
45	336.5	2.2	3411	1	GNWVYP	genome polyprotein

ALIGNMENTS

```
RESULT 1
GNWVJ8
genome polyprotein - hepatitis C virus (strain HC-J8)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
   protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A40250; PQ0397; PQ0559
R;Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Iizuka, H.; Tanaka, T.; Fukuda, S.;
  Virology 188, 331-341, 1992
A;Title: Full-length sequence of a hepatitis C virus genome having poor homology to repo
A;Reference number: A40250; MUID:92230232; PMID:1314459
A;Accession: A40250
A;Molecule type: genomic RNA
A;Residues: 1-3013 <OKA>
A;Cross-references: UNIPROT:P26661; GB:D10988; GB:D01221; NID:g221608; PIDN:BAA01761.1;l
R;Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.L.
  J. Gen. Virol. 73, 1131-1141, 1992
A;Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e
A;Reference number: PQ0393; MUID:92268871; PMID:1316939
A;Accession: PQ0397
A;Molecule type: genomic RNA
A;Residues: 2678-2754 <CHA>
A;Cross-references: DDBJ:D10134
A;Experimental source: isolate E-b12
R;Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno,
  Biochem. Biophys. Res. Commun. 181, 279-285, 1991
A;Title: Distribution of plural HCV types in Japan.
A;Reference number: PQ0554; MUID:92068204; PMID:1720309
A;Accession: PQ0559
A;Molecule type: mRNA
A;Residues: 2678-2729 <KAT>
A;Cross-references: GB:D10562; GB:D90518; NID:g221523; PIDN:BAA01418.1; PID:g221524
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural]
F;1-115/Product: envelope protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-733/Product: nonstructural protein NS1 #status predicted <NS1>
F;734-1010/Product: nonstructural protein NS2 #status predicted <NS2>
F;1011-1619/Product: nonstructural protein NS3 #status predicted <NS3>
F;1234-1241/Region: hepacivirin #status predicted <KAT>
F;1316-1321/Region: nucleotide-binding motif A (P-loop)
F;1320-1323/Region: nucleotide-binding motif B
F;1320-1323/Region: DEXH motif
F;1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
F;2018-3033/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,233,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2018,23

Query Match                 21.3%;  Score 3273.5;  DB 1;  Length 3033;
Best Local Similarity       21.9%;  Pred. No. 3.4e-195;
Matches  980;  Conservative  436;  Mismatches 1185;  Indels  679;  Gaps  106;
```

us-10-009-002-2.rpr

```
Qy    8  TSPVPAPRTRKQKQTQASYPVSIK-------------------TSVERGQRAKRKVQRDARPR   51
                  :  :    :  ||  |: |        |||
Db    3  TNPKPQRRTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGRPRLGVRATRKTSERSQPR    59

Qy   52  NYKIAGIHDGLQTLAQAALPAH----------------------GWGRQDPRHKSRNL     87
           |:  |  |||                                 || || ||||||
Db   60  GRRQPIPKDRRSTGKSWGKPGYPWPLYGNEGCGWAGWLLSPRGSRPTWGPTDPRHRSRNL   119

Qy   88  GILLDYPLGWIGDVTTHTPLVGPLVAGAVRPVCQIVRLLEDGVNWATG---WFGVHLFV    144
           |::|  | |||| |  :  ||:|| ||  |  |  ||||||:| ||      |||: |:
Db  120  GRVIDTITCGFADLMGYIPVVGAPV-GGVARALAHGVRVLEDGINYATGNLPGCSFSIFL    178

Qy  145  VCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVICADE-----   198
          | ||  |  | |:|  ::  | | :|:| ::::: |  |:  || | || | |:
Db  179  LALLSCVVPVSAVEVRN-ISSSYYATNDCSNNSITTWQLTDAVLHLPGCVPCENDNGTLH   237

Qy  199  CWPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV----   254
          || :||::||: | | :::: |::||:| |:||:||| |  | | | |::  ::
Db  238  CWIQVTPNVAVKHRGALTRS-LRTHVDMIVMAATACSALYVGDVCGAVMILSQAFMVSPQ   296

Qy  255  RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTKLASQVPYAIATMFS   314
          ||    | :|  |   :      :|:    :|:   | : |:  |:  :  ||: : |:
Db  297  RH-------NFTQECNCSIYQGHITGHRMAWDMMLSWNSPTLTMILAYAARVPELVLEIIF   349

Qy  315  SVHYLAVGALIIYASRGKWYQLLLALMLY--IEAT---SGNPIRVPTGCSIAEFCSPLMI   369
          | :  |||| :: |  |   ::| | |:    |:      | ||     || |   | :
Db  350  GGHWGVVPFGLAYPSMQGAMAKVIAILLLVAGVPATTYSSGQ----EAGRTVAGFAGLFTT   405

Qy  370  PCP----------------------CHSYLSENVSEVICYSPKWT----------RPI-   395
          ||                                   |
Db  406  GAKQNLYLINTNGSWHINRTALNCNDSLQTGFLASLFYTHKFNSSGCPERLSSCRGLDDF   465

Qy  396  ----TLEYNNSIS----------W-YP---------YTIPGARGCM-------------   417
              :||:                 ||          | |||| |
Db  466  RIGWGTILEYETNVTNDGDMRPYCWHYPRPCGIVPARTVCGPVYCFTPSPVVVGTTDKQG   525

Qy  418  ---------------VKFKNNT--------WGC----------CRIR-----NVPSYC    437
                         || :           | |           || |     || | |
Db  526  VPTYTWGENETDVFLLNSTRPPRGAWFGCTWMNGTGFTKTCGAPPCRIRKDYNSTIDLLC   585

Qy  438  TMGTDAVWNDTRNTYEACGVTPWLT----------TAWHNGSALKLAILQ---YPGSKEMF  485
             |:||   |  |     |  |||           ||::|||:  ||:|     | |  :
Db  586  P--TDCFRKHPDATYLKCGAGPWLTPRCLVDYPYRLWHYPCTVNFTIFKARMYVGGVE--   641

Qy  486  KPHNWMSGHLYP-----------EGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGS   534
           ||: : |:               |                |     |
Db  642  --HRFSAACNFTRGDRCRLEDRDRGQQSPLLH-------ST---TEWAVLPCS-------   682

Qy  535  WLQVPQGFYSDVKDLATGLIITKDKAWKNYQVLYSATGALSLTGVTTKAVVLLLGLCGSK   594
                 |||: :      ||: :::   || :| |:: :  | |    || |:::| :  |
Db  683  -------FSDLPALSTGLLHLHQNIVDVQYLYGLSPALTRYIVKWEWWILLFLLLADAR-   734

Qy  595  YLILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYLRCR   654
           : :|  | ||                |::  | | :::   | || ||: |: |
Db  735  --ICACLWMLI--------------------ILGQAEAALEKLII---------------  757

Qy  655  LRYAALLGFVPMAAGLPLTFFVAAAAQPDYDWVRLLVAGLVLWAGRNRGHRIALLVGP    714
           : :|: | :|:||:  ||||: : |:       |  ||:  |   :     : :|
Db  758  -----------LHSASAASANGPLWFFIFPTAA-------WYLKGRVVPVATYS-----  797

Qy  715  WPLVALLTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLW-  773
           | :|||| ||:  : |   :: |:|| | ::  ||::||     ::|||  ::||:::
Db  798  WSF--LLLVLALPQQAYALDAAEQGELGLAILVIISIFTLTPAYKILLSRSVWWLSTMLV  855

Qy  774  QRW-------ENWFW-NVTLRPERFFLV--LVCFPGATYDALVTFCVCHV           813
           :         :   |  :|:||    :    | ||| |: | | ||
Db  856  LAEAQIQQWVPPLEVRGGRDGIIWVAVILHPRLVFEVTKWLLAILGPAY-----       904

Qy  814  ALLCLTSSAASFFGTDSRVRAHRML--VRLGKCHAWYSHYVLKFFLLVFGENGVFFYKHL  871
           :|  |:|:|:|| |||  ||||: |   | |     | :|| :|   :| ::|  : :||
Db  905  -ILLKASLLRIPYF------VRAHALLRVCTLVK-HLAGARYIQMLLIITIGRWTGTYIYDHL  957
```

```
Qy  872  HGDVLPNDFASKLPLQ----------------------EP--FFPPEGKARVYRNEGRRLACGDTVDG  915
           || |:|:|||| ||                       |    |||||| ||| | |||:| |||:| |
Db  958  S-------------------------PLSTWAAQGLRDLAIAVEPVVFSPMEKKVIWGAE--TVACGDILHG 1003

Qy  916  LPVVARLGDLVFAGLAMPPD-----GWAITAPPTLQCLSERGTLSAMAVVMTGIDPRTWT   970
           ||::|||| :|  ||||          |:| |||  |  : |||: |:|  |||| ::|
Db 1004  LPVSARLGREVLLG---PADGYTSKGWKLLAPITAYTQQTRGLLGAIVUSLTGRDKNEQA  1060

Qy  971  GTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDIYQPPCG  1030
           | ::|| || || :|:: |:|| :| ||: |: || || | |:|| :: : |:::| ||
Db 1061  GQVQVLSSVTQTFLGTSISGVLWTVYHGAGNKTLAGPKGPVTQMYTSAEGDLVGWPSPPG  1120

Qy 1031  AGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHV  1090
            |||| |:||    | |||||||||:|||| |   |  ||| : |:||||||:|| ||||
Db 1121  TKSLDPCTCGAVDLYLVTRNADVIPVRRKDDRRGALLSPRPLSTLKGSSGGPVLCSRGHA  1180

Qy 1091  IGMFTA------ARNSGGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILIAPTG 1147
            :|:||        : :| |||||: |:  | |::| ||:| ||  |  :|  || |||: |||
Db 1181  VGLFRAAVCARGVAKSIDFIPVESLDVATRTPSFSDNS---TPPAVPQSYQVGVLHAPTG  1237

Qy 1148  SGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCNTNGASLTY  1207
           ||||||:|  |  | | ||||||||| |   : ||| ||| ||| |:|| :| || | ||
Db 1238  SGKSTKVPAAYASQGYKVLVLNPSVAATLGFGAYMSKAHGINPNIRTGVRTVTTGDSITY  1297

Qy 1208  STYGMYLT-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPG  1265
           |||| |   | |    ||:|||||||: ||||:|||| ||| | : ||||||||||||||
Db 1298  STYGKFIADGGCAAGAYDIICDECHSVDATTILGIGTVLDQAETAGVRLVVLATATPPG   1357

Qy 1266  VIPTPHANITEIQLTDEGTIPFFHGKKLKEENLKKGRHLIFEATKHGCDELANELARKGIT 1325
           | || |:| |||:| :|| ||:   |  ||||||||||||| ::|  ||||| | |: |
Db 1358  TVTTPHSNIEEVALGHEGEIPFYGKAIPLAFIKGGRHLIFCHSKKKCDELAAALRGMGVN  1417

Qy 1326  AVSYIRGCDISKIP-EGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMG  1384
           |: | ||  :| || :||||||||||| ||||||||||:||:: |   | ||||||| :
Db 1418  AVAYYRGLDVSVIPTQGDVVVVATDALMTGYTGDFDSVIDCNVAVSQIVDFSLDPTFTIT  1477

Qy 1385  VRVCGVSAIVKGQRRGRTGRGRAGIYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSST  1444
             | |:|:: ||||||||||| || | :  | | ||| : : : :  |||:||| |:
Db 1478  TQTVPQDAVSRSQRGRTGRGRLGVYRYVSSGERPSGMFDSVVLCECYDAGAAWYELTPA  1537

Qy 1445  EAQTILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLQL  1503
           |  | |  | ||||||  | :| :|| :   :|: ||||| | |:| ||: | ||:|  :
Db 1538  ETTVRLRAYFNTPGLPVCQDHLEFWEAVFTGLTHIDAHFLSQTKQGGENFAYLTAYQATV  1597

Qy 1504  CHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADAC----------PGPEP------   1545
           | :  ||   | |                :|  ||
Db 1598  CAR----AKAPPSWD-------VMWK------CLTRLKPTLTGPTPLLYRLGAVTN     1637

Qy 1546  ------SEVTRY----QMCFTEVNTSGTAALAVGVAMAYLAIDTFGATCVRRCWSITS-  1595
                     :  :|  |    : | :: |:|  | ||  :
Db 1638  EVTLTHPVTKYIATCMQADLEIMTS-SWVLAGGVLAAVAAYCLAT----------GCISIIGR  1689

Qy 1596  --VPTGATVAPVVDBE---------EIVEECASFIPL-EAMVAAIDKLKSTI------TTTSP  1640
              :  |   |::|: :        | :||| |||    |:  |   |:|
Db 1690  LHLNDRVVVAP--DKEILYEAFDEMEECASKAALIEEGQRMAEMLKSKIQGLLQQATRQA  1747

Qy 1641  FTLETALE---KLNTFLGPHAATILAIIEYCCGLVTLPDNPFASCVPAFIAGITTPLPH  1696
            |:|  |     |::::||  | || |||| || |||||||  :|| ||:||::|:||
Db 1748  QDIQPAIQSSWPKLEQFWAKHMWNFISGIQYLAGLSTLPGNPAVASMMAFSAALTSPLPT  1807

Qy 1697  KIKMFLSLFGGAIASKLTDARGALAFMAGAAGTALGTWTSVG---FVFDMLGGYAAASS  1753
           : ::  ::  | :||  : || |: :: |: | |:|  : |     :::::| ||| |
Db 1808  STTILLNIMGGWLASQIAPPAGATGFVVSGLVGAAVG---SIGLGKILVDVGYIAGYGAGIS  1864

Qy 1754  TACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTTAGPD--HWPNRLL  1811
            | ::|| :|||:||::|::  ::: | | || |||: | |:| | | :     || ||:
Db 1865  GALVAFKIMSGEKPTVEDVVNLLPAILSPGALVVGVICAAILRRHVGQGEAVQMNRLI   1924

Qy 1812  TMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTEDDCGLI---AW  1868
            : | ||   |::  ::  | |::   :: |  |:  :: || ||| ||| || :   ||
Db 1925  AFASRGNHVAPTHYVVESDASQRVTQVLSSLT----ITSLLRRLHAWITEDCPVPCSGSW  1980

Qy 1869  GLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCCQKGYKGPWIGSGMLQARCPCGAE 1928
```

Db   2995 VGAGGGDIYHSVSHARPRLL------LLCLLLLSVGVGI 3027

RESULT 2
JQ1303
genome polyprotein - hepatitis C virus (isolate HC-J6)
N;Contains:capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
  protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
A;Accession: JQ1303
R;Okamoto, H.; Okada, S.; Sugiyama, Y.; Kurai, K.; Iizuka, H.; Machida, A.; Miyakawa, Y.;
  J. Gen. Virol. 72, 2697-2704, 1991
A;Title: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a hume
A;Reference number: JQ1303; MUID:92044440; PMID:1658196
A;Accession: JQ1303
A;Molecule type: genomic RNA
A;Residues: 1-3033 <OKA>
A;Cross-references: UNIPROT:P26660; GB:D00944; NID:g221650; PIDN:BAA00792.1; PID:g221651
A;Experimental source: isolate HC-J6 from a Japanese individual
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; P-loop; polyprotein; serine proteinase; transme
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-733/Product: nonstructural protein NS1 #status predicted <NS1>
F;734-1010/Product: nonstructural protein NS2 #status predicted <NS2>
F;1011-1619/Product: hepacivirin #status predicted <NS3>
F;1316-1321/Region: nucleotide-binding motif B
F;1320-1323/Region: DEXH motif
F;1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
F;2018-3033/Product: nonstructural protein NS5 #status predicted <NO5>
F;196,209,234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2038,281

Query Match   21.2%;   Score 3262;   DB 1;   Length 3033;
Best Local Similarity 30.1%;   Pred. No. 1.8e-194;
Matches 980;   Conservative 432;   Mismatches 1205;   Indels 640;   Gaps 112;

Qy   8 TSPVPAPRTRKNKQTQASYPVSIK------------------TSVERGQRAKRKVQRDARPR 51
              :   :   : ::   :   :||
Db   3 TNPKPQRTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR 59

Qy   52 NYKIAGIHDGLQTLAQAALPAH------------------------GWGRQDPRHKSRNL 87
             |: ||| || :: | | | |:| |||: |
Db   60 GRRQPIPKDRRSTGKSWGKPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPNDPRHSRNV 119

Qy   88 GILLDYPLGWIGDVTHTPLVG-PLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLF 143
             || | | ||| : | ||| || :| | : || |: ||| :|
Db   120 GKVIDTLTCGFADLMGYIPVVGAPL--GGVARALAHGVRVLEDGVNFATGNLPGCSFSIF 177

Qy   144 VVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC------AD 197
             :|| |: ||| ||| ||| | | || | | | | | | | |
Db   178 LLALLSCITTPVSAAEVKNISTG-YMVTNDCTNDSITWQLQAAVLHVPGCVPCEKVGNTS 236

Qy   198 ECWVPANPYIS--HPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV- 254
             : ||| | : || | | | | | | : | | |
Db   237 RCWIPVSPNVAVVQPGALT--QGLRTHIDMVVMSATLCSALYVGDLCGGVMLAAQMFIV 293

Qy   255 ---RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTKLASOVPYAIAT 311
                 : | |: : | || | | | | || | || : | :   :| |
Db   294 SPQHHWFVQ----DCNCSIYPGTITGHRMAWDMMNWSPTATMILAYAMRVPEVIID 346

Qy   312 MFSSVHYLAVGALIYYASRGKWYQLLLALML-------- 342
             :|:: | | :: | | || : | :|
Db   347 IIGGAHWGVMFGLAYFSMQGAWAKVVVLLLAAGVDAQTHTVGGSTAHNARTLTGMFSLG 406

Qy   343 ---------------------------YIEAT------SGNPIRVPTGCSIAEF- 363
                                         || | || ||: | | |||
Db   407 ARQKIQLINTNGSWHINRTALNCNDSLHTGFLASLFYTHSFNSSGCPERMSACRSIEAFR 466

Qy   364 -----------------CSPLMIPCPCHSY----------LSENVSEVICYSPKWTRPITLEYNNS 402
                              | | || |:| |:| :: | | | | ||: |

Db   1981 LQDIWDWVCSILTDFKNWLSS--KLLPKMPGIPFISCQKGYKGVWAGTGVMTTRCPCGAN 2038
Qy   1929 LIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARP-DPTDWTSLVVNYGVRDYCKY 1987

Db   2039 ISGHVRMGTMKI-TGPKTCLNLWQGTFPINCYTEGPCVPKPPPNYKTAIWRVAASEYVEV 2097
Qy   1988 EKMGDHIFVTAVSSPNVCFTQVP---PTLRAAVADGVQVQCYLGEPKTPWTTSACCYGP 2044

Db   2098 TQHGFSFYVTGLTSDNL---KVPCQVPAPEFFSWVDGVQIHRFAPVP-----------GP 2143
Qy   2045 DGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSD---EAAVSALVPKQEL 2101

Db   2144 FFRD---EVTFTVGLNSFVVGSQLPCDPEPDTEVLASMLTDPSHITAEAAARRLARGSPP 2200
Qy   2102 RRTNQLLEAISA---------------------------GVDTTKLPA----------- 2122

Db   2201 SQASSSASQLSAPSLKATCTTHKTAYDCDMVDANLFMGGDVTRIESDSKVIVLDSLDSMT 2260
Qy   2123 -------PSI-EEVVVRKRQPRARTGSLTLPPPPRSVPGVSCPESLQRSDPLEGPSNLPP 2174

Db   2261 -----ALPPWAR--------ALPPWAR 2292
Qy   2175 SPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKKE----------- 2221

Db   2293 YNPVLIETWKRP------GYEEPTVLGCALPPT-----PQTPVPPPRRRAKVLTQDNVEGV 2343
Qy   2222 ----------------------VSEWSDESWSTATTASSYVTGPPYPKIRGK 2251

Db   2344 LREMADKVLSPLQDNNDSGHSTGADTGGDIVQQPSDE--TAASEAGSLSSMPPLEGEPGD 2401
Qy   2252 DSTQ------SAPAKR------------PTKKKLGKSEFSCCSMSYTWTD-VISFKTASKV 2292

Db   2402 PDLEFEPPVGSAPPSEGECEVIDSDSKSWSTVSDQEDSVICCSMSYSWTGALITPCCGPEEE 2461
Qy   2293 LSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQVRLAKEKASKVV 2352

Db   2462 KLPINPLSNSLMRFHNKVYSTTSRSASLRAKKVTFDRVQVLDAHYDSVLQDVKRAASKVS 2521
Qy   2353 GVMWDYDEVAAHTPSKSAKSHITGLRCTDVRSGAARKAVLDLQKCVEAGEIPSHYR-QTV 2411

Db   2522 ARLLTVEEAACALTPPHSAKSRY-GFGAKEVRS-LSRRAVNHIRSVWEDLLEDQHTPIDTT 2579
Qy   2412 IVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGGQVAPDVVKAVMGDAYGF-VDP 2470

Db   2580 IMAKNEVFCIDPTKGGKKPARLIVYPDLGVRVCEKMALYDIAQKLPKAIMGPSYGFQYSP 2639
Qy   2471 RTRVKRLLSMW---SPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQHRAGIHTIA 2528

Db   2640 AERVDFLLKAWGSKKDPMGFSYDTRCFDSTVTERDIRTEESIYQACSLPQEARTVIHSLT 2699
Qy   2529 RQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMKNPRFLICG 2588

Db   2700 ERLYVGGPMTNSKGQSCGYRRCRASGVFTTSMGNTMTCYIKALAACKAAGIVDPVMLVCG 2759
Qy   2589 DDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSNVTSGITKS 2648

Db   2760 DDLVVISESQGNEEDERNLRAFTEAMTRYSAPPGDLPRPEYDLELITSCSSNVSVALDSR 2819
Qy   2649 GKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHFMEQMLFED 2708

Db   2820 GRRRYFLTRDPTTPITRAAWETVRHSPVNSWLGNIIQYAPTIWVRMVIMTHFFSILLAQD 2879
Qy   2709 KLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNABIILRVSQSLTDMTMPPL 2768

Db   2880 TLNQNLNFEMYGAVYSVNPLDLPAIIERLHGLEAFSLHTYSPHELSVAATLRKLGAPPL 2939
Qy   2769 RAWRKKARAVLASAKRGGAHAKLARPLL-WHATSR---PLPDLDKTSVARYTTFNYCD 2823

Db   2940 RAWKSRARAVRASLIAQGARAAICGRYLFNWAVKTKLKLTPLPEASRLDLSGWFT----- 2994
Qy   2824 VYSPEGDVFITPQRRIQKFLVKYLAVIVFALGLIAVGLAI 2863

```
Qy 1401 RTGRGRAGIYYYVDGSCTPSGMVPECNIVEAPDAAKAWYGLSSTEAQTILDYRTQPGLP 1460
        ||||||| |||| |  |  | ::  :  :   ::  ||   :          ::||||
Db 1494 RTCRGRLGIYRYVVSTGERASGMFDSVVLCECYDAGAAWYELTPAETTVRLRAYFNTPGLP 1553

Qy 1461 AIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLQLCHQYGYAAPNDAPRWQ 1519
        . | :|:||| ::  ||:|   |.  : :: .||  ||| :| | .|| |||    .
Db 1554 VCQDHLEFWEAVFGLTHIDAHFLSQTKQSGENFAYLTAYQATVCAR----AKAPPSWD 1609

Qy 1520 GARLGKKPCGVLWRLDGADAC---------PGPEP----------SEVTRY----Q 1552
                     :|                 |                 ::|
Db 1610 ---------VMWK-----CLTRLKPTLVGPTPLLYRLGSVTNEVTLTHPVTKYIATCM 1653

Qy 1553 MCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEEI- 1611
         | :|: ::  :|| | |::|    || .  | ||          | |     :|:||
Db 1654 QADLEVMTS-TWVLAGGVLAAVAAYCLAT-GCVCI------IGRLHVNQRAVVAP--DKEVLY 1706

Qy 1612 ------VEECASFIPL-EAMVAAIDKLKSTI--------TTTSPFFLETALE---KLNTFLG 1655
              |||||:  | .||:|| |:|:|:|          |::| :: ::|:|
Db 1707 EAFDEMEECASRAALIEEGQRIAEMLKSKIQGLLQQASKQAQDIQPAVQASWPKVEQFWA 1766

Qy 1656 PHAATILAIIEYCCGLVTLPDNPFASCVFAFIAGITTTPLPHKIKMFLSLFGGAIASKLTD 1715
         |   |  :|:|  ||::|| ||  :   |: |   : |:|:|  ::  ::||  |::  :
Db 1767 KHMWNFISGIQYLAGLSTLPGNPAVASMMAFSAALTSPLSTSTTILLNILGGWLASQIAP 1826

Qy 1716 ARGALAPMMAGAAGTALGTWTSVGF---VFDMLGGYAAAASSTACLTFKCLMGEWPTMDQL 1772
           | || :.| | |: |   :.|     . ||.|||| : |: |  || .|.||:| |:
Db 1827 PAGATGFVVSGLVGAAVG---SIGLGKVLVDILAGYGAGISGALVAFKIMSGEKPSMEDV 1883

Qy 1773 AGLVYSAFNPAAGVGVLSACAMFALTTAGPD---HWPNRLLTMLARSNTVCNEYFIAT 1828
        . |:  : ||  |:|:|.|    |:   ||     ||||||| :|  | ||::|  | :|
Db 1884 VNLLPGILSPGALVVGVI--CAAILRRHVGPGEGAVQWMNRLLIAFASRGNHVAPTHVVTE 1941

Qy 1829 RDIRRKILGILEASTPWSVISACIRWLHTPTEDDCGLI---AWGLEIWQYVCNFFVICFN 1885
        .| ||::|| || :  |||:|: :|  | .  |:  |     :|.:  |:|| :|  : ||
Db 1942 SDASQRVTQLLGSLT----ITSLLRRLHNWITEDCPIPCSGGSWLRDVWDWVCTILTDFKN 1997

Qy 1886 VLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPR 1945
         |::|:|| | ||| |: ||| | ||.| |:| :   ||||| ::| | || :   :|
Db 1998 WLTS--KLFPKMPGLPFISCQCKYKGVWAGTGIMTTRCPCGANISGNVRLGSMRI-TGPK 2054

Qy 1946 TCSNYWRGAVPVNARLCGSARPDT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV 2004
        || | |:|:.| | ||||:.|||  ::  ::.|| ..::  |  .|    ||.:  :|:
Db 2055 TCMNIWQGTFPINCYTEGQCVPKPAPNFKIATWRVAASEYAEVTQHGSYHYITTGLTTDNL 2114

Qy 2005 CFTQVP---PTLRAAVAVDGVQVQVQCYLGEPKTPWTTS-ACCYG----------PD 2045
          :|     | .. |: ||||: .:    : |: |   | .|
Db 2115 ---KVPCQLPSPEFFSWVDGVQIHRFAPIPKPFFRDEVSFCVGLNSFVVGSQLPCDPEPD 2171

Qy 2046 GKGKTVKL--PPRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQELRR 2103
        | ||||      |:||.|   || :  :  ||||  . |::||:. |: ||  :::: |
Db 2172 TDVLTSMLTDPSHITAETAARRL---------ARGSPPSEASSSASQLSAPSLRA 2217

Qy 2104 T------NQLLEA-ISAGVDTTKLPAPSIEEVVVRKRQFRARTGSLTLPPPRSVPG 2153
                .:||:| ::|| |   : :  .: .| ||:
Db 2218 TCTTHGKAYDVDMVDANLFMGGDVTRIESES--KVVV------------------ 2252

Qy 2154 VSCPESLQRSDPL-EGPSNLPPS-----------PPVLQL-AMP---MPLLGA---GEC 2193
        ||| :.| : ||  :||:||||             |:| | | |   :||   |
Db 2253 ------LDSLDPMVEERSDLEPSIPSEYMLPKKRFPPALPAWARPDYNPPLVESWKRPDY 2306

Qy 2194 NPFTAIGCAMTETGGGPDLPSYPPKKE-----VSEWS---------------DES 2229
         | :|.||:: . |.| || |: |||      |||:
Db 2307 QPATVAGCALPP----PKKTPTPPPRRRRTVGLSESSIADALQQLAIKSFGQPPPSGDSG 2362

Qy 2230 WST-ATTASSYVTGPP---------------YPKIRGKDSTQSAPAKRPTKKKL---- 2267
        .|| |  |||:|||                 ||::||::|.  |||:|| |||:|
Db 2363 LSTGADADSGGRRTPPDELALSETGSISSMPPLEGEPGD---PDLEPEQVELQPPPQGGV 2419

Qy 2268 -----GKSEFS---------------CSMSYTWTD-VISFKTASKVLSATRAITSGFLKQRSLV 2310
             || |:                ||||| |.| || |:|||||:||| |||:| ||::|:|
Db 2420 VTPGSGSGSWSTCSEEDDSVVCCSMSYSWTGALITPCSPEEEKLPINPLSNSLLRYHNKV 2479
```

```
Db 467 VGWGALQYEDNVTNPEDMRPYCWHYPPRQCGVVSASSVCGPVVCFTPS---PVVVGTTDR 523
       : | |:| |  |||| :| | |     :|  |   | |||| | |     ||| |.|:
Qy 403 ISWYPYTIPGARGCMVKFKNNT--------CRIR---NVPSY 436

Db 524 LGAPTYTW-GENETDVFLLNSTRPPQGSWFGCTWMNSTGYIKTCGAPPCRIRADFNASMD 582
       | | | |  |:|.|:|   .:| |. |.| | |||||.| :||| | ||||||| |::|
Qy 437 CTMGTDAVWNDTRNTYEACGVTPWLT--------TAWHNGSALKLAILQ--YPGSKEM 484

Db 583 LLCPTDCFRKHPDTTYIKCGSGPWLTPRCLIDYPYRLWHYPCTVNYTIFKIRMYVGGVE- 641
       :|  || |::||:|| ||| | || | ||:||||||||||||||| ||||||||||||
Qy 485 FKPHNVMSGHLYFEGSDTPIVFYDPVNSTLLP----PERWARLPGTPPVVRGSWLQVPQ 540

Db 642 ---HRLTAACNFTRGDRCNL---EDRDRSQLSPLLHSTEWAILPCT 682
          |:|:||||||||||||    |:||||| |:||| ||||:||||
Qy 541 GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTCGVTTKAVVLILGLCGSKYLILAY 600

Db 683 --YSDLPALSTGLLHLHQNIVDVQFMYGLSPALTKYIVRWEWNVVLLFLLLADAR--VCAC 738
         |||: | |:|:||||:|:|||||||||||||:|||:|||||| |||:|||||   |||
Qy 601 LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYLRCRLRYAAL 660

Db 739 -------LGQAEAALEKLLVVLHAA-----------SAASC 766
              || || |:|:||:|||||            || |
Qy 661 LGFVPMAAGLPLTFFVAAAAAAQCPDYDWMVRLLVAGLVLWAGRNRGHRIALLVGPWPLVAL 720

Db 767 NGFLYF-----VIFFVAA----------WIKGRVVPLATYS--------LTGLWSFG-- 801
       ||||||     |||||            ||||||:||||:|        |||||||
Qy 721 LTLLHLVTPASAFDTELIGGLTIPPVVALVMSRFGFFAHLLPRCALVNSYLWQRW---- 776

Db 802 LLLALPQQAYAYADASVHGQIGAALLVLITLFTLTPGYKTLLSR------FLW--WLCYL 853
       ||||||||||||||||||||||||||||||||||||||||||||      |||  ||||
Qy 777 -----ENWFWNVTLRPER---FFLVLVCFPGATVDALVTFCVCHVALLCLTSSAASF 825

Db 854 LTLAEAMVQEWAPPMQVRGGRDGIIWAVAIFCPGVVFDITKW---------LLAVLGPAYLL 906
       |||||||||||||||||||||||||||||||||||||||||          |||||||||
Qy 826 VRAHRML-----VRLGKCHAWYSHYVLKFFLLVFGE-NGVFFYKHL--HG 873

Db 907 KGALTRVPYFVRAHALLRMCTMVR----HLAGGRYV-QMVLLALGRWTGTYIYDHLTPMS 961
       :|||||||||||||||||||||     |||||||| |||||||||||||||||||||||
Qy 874 DVLPNDFASKLPLQEP---FFPPEGKARVYRNEGRRLACGDTVDGLPVVARLGDLVPAGLA 931

Db 962 DWAANGLRDLAVAVEPIIFSPMEKKVIVWGAE--TAACGDILHGLPVSARLGREVLLG-- 1017
       |||||||||||||||||||||||||||||||   |||||||||||||||||||||||
Qy 932 MPPD--------GWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSMGF 986

Db 1018 -PADGYTSKGWSLLAPITAYAQQTRGLLGTIVVSMTGRDKTEQAGEIQVLSTVTQSFLGT 1076
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 987 VCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCSCGETKGYL 1046

Db 1077 TISGVLWTVYHGAGNKTLAGSRGPVTQMYSSAEGDLVGWPSPPGTKSSLEPCTCGAVDLYL 1136
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1047 VTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARRNSGG---S 1103

Db 1137 VTRNADVIPARRRGDKRGALLLSPRPLSTLKGSSGSGGPVLCPRGHAVGVFRAAVCSRGVAKS 1196
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1104 VSQIRVRPLVCAGYHPQVTAHATLDTKPTVPNEYSVQLLIAPTGSGKSTKLPLSYMQEKY 1163

Db 1197 IDFIPVETLDIVTRSPTFSDNS---TPPAVPQTYQVGYLHAPTGSGKSTKVPVAYAAQGY 1253
       |||||||||||||||||||||    |||||||||||||||||||||||||||||||||||
Qy 1164 EVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYLT-GACSRN- 1221

Db 1254 KVLVLNPSVAATLGFGAYLSKAHGINPNIRTGVRTVTTGAPITYSTYGKFLADGGCAGGA 1313
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1222 YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLVLATATPPGVIPTPHANITEIQLTD 1281

Db 1314 YDIIICDECHAVDSTTLGIGTVLDQAETAGVRLTVLATATPPGSVTTPHPNIEEVALGQ 1373
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1282 EGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELARKGITAVSYYRGCDISKIP-E 1340

Db 1374 EGEIPFYGRAIPLSYIKGGRHLIFCHSKKKCDELAAALRGMGLNAVAYYRGLDVSVIPTQ 1433
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1341 GDCVVVATDALCTGYTGDFDSVYDCSLMVEGTHCVDLLDPTFTMGVRVCGVSAIVKGQRRG 1400

Db 1434 GDVVVVATDALMTGFTGDFDSVIDCNVAVTQVVDFSLDPTFTITTQTVPQDAVSRSQRRG 1493
```

```
Query Match        21.1%;    Score 3237;    DB 1;    Length 3010;
Best Local Similarity 30.1%;  Pred. No. 6.4e-193;
Matches    980;  Conservative 437;  Mismatches 1192;  Indels 646;  Gaps 115;

Qy     4  ISTQTSVPAPRTRKNKQTQASYPVSIK-----------------TSVERGQRAKRKVQRD  47
          :||:|:  :|| |
Db     1  MSTNGKP------QRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTWER  55

Qy    48  ARPRN------------------YKIAGIHDGLQTLAQAALP---AHGWGRQDPRH  82
          : ||
Db    56  SQPRGRRQPIPKARQPEGRAWAQPGYPWPLYG-NEGLGWAGWLVSPRGSRPNWGPTDPRR  114

Qy    83  KSRNLGILLDYPLGWIGDVTTHTPLVG-PLVAGAVVRPVCQIVRLLEDGVNWATG---WF  138
          :||||| :|  |    |:  |||| ||   :|| |   |  |||||||||| ||
Db   115  RSRNLGKVIDTLTCGFADLMGYIPLVGAPL--GGVARALAHGVRVLEDGVNVATGNLPGC  172

Qy   139  GVHLFVVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVICAD  197
          . : |:|| :| | :  |.|:|   : |: | :|| :|: :|:        ||| :|
Db   173  SFSIFLLALLSCLTIPASAYEVHNV-SGIYHVTNDCSNSSIVYEAADMIMHTPGCVPCVR  231

Qy   198  E------CWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDW  252
                :|| |    :   | ||    :    |  :   ||||   ::  ||| ||||:|
Db   232  ENNSSRCWVALTPTLAARNNSVPT-ATIRRHVDLLVGAAAFCSAMYVGDLCGSVFLVSQL  290

Qy   253  LV--RHWLIHIDLNETGTCVLEVPTGIDPGFL--GPIGW---MAGKVEAVIFLTKLASQ  304
          |    :|||||| |||| || ||| | |||    ||:|    |||:|||| ||| |||:
Db   291  --------------------------CSIYPGHVTGHRMAWDMMMNWSPTTALVVSQLL-R  339

Qy   305  VPYAIATMPSSVHYLAVGALIYASRGKWYQLLLALMLY-----------  343
          :| ||  | :| :|: |||: :| |||||| | |: |
Db   340  IPQAVVDMVGGAHWGVLAGLAYYSMVGNWAKVLIVMLLFAGVDGSTIVSGGTVARTHSL  399

Qy   344  --------------------------IEAT---------------
                                    ||:|
Db   400  ASLFTQGASQKIQLINTNGSWHINRTALNCNDSLQTGFLASLFYAHRFNASGCPERMASC  459

Qy   358  CSIAEF----CSPL-------------------MIPCPCHSYLSENV-SEVICYSPKWTRPITL  397
          ||:| |    | ||                   ||||  | |   :| ||::||||||| |:|
Db   460  RSIDKFDQGWGPITYTEADIQDQRPVCWHYAPRPCGIVPASQVCGPVVCFTPS---PVVV  516

Qy   398  EYNN-----SISWYPYTIPGARGCMVKFKNNT----------WGCCRIRNVPSY-------CT  438
          |:      |::|:   :| |  | ||||:|            ||| ||:| |:       ||
Db   517  GTTDRFGAPTYSW-------GENETDVLILNNTRPPQGNWFGCTWMNSTGFTKTCGGPPCN  570

Qy   439  MG--------TDAVWNDTRNTYEACGVTPWLT----------TAWHNGSALKLAILQ--  477
          :|        || :  | || |||| |||||          | |||||:||:||||
Db   571  IGGGGNNTLVCPTDCFRKHPEATYTKCGSGPWLTPRCMVDYPYRLWHYPCTVNFTIFKVR  630

Qy   478  -YPGSKE--MFKPHNWMSG-HLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRG  533
            |||:   :| || |  | ||::| || |:   : |::  || :|:| |   || | |
Db   631  MYVGGVEHRLNAACNWTRGERCDLEDRDR-------SELSPLLLSTTEWQILPCS------  678

Qy   534  SWLQVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGS  593
          :| :|| | : || |||  |: || |   ||:|  |:| |||| : ::|::||: |||||
Db   679  --------FTGLPALSTGLIHLHQNVVDVQYLYGIGSAVVSPAIKWEYILLLFLLLADA  729

Qy   594  KYLILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRC  653
          : |  |  :::|: |                    :|   :|| |||:||:  :|  |: |
Db   730  RVCACLWMMLL-------------------IAQAEAALENLVVFNAASV--  759

Qy   654  RLRYAALLGFVPMAAGLPLTFVAAAAQPDYDWVR-LLVAGLVLWAGRNRGHRIALLV  712
          |  :   :|:  | |||| |:  | :      |   |::||:| | | |  : |: |
Db   760  ----AGMHGTLSF-----LVFFCAA------WYIKGRLVPGAAY------ALY  791

Qy   713  GPWPLVALLTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYL  772
          | |||: |||:|||:||| ||  :    ||:|| |  || || || | ||  |||||:|
Db   792  GVWPL--LLLLALPPRAYAMDREMAASCGGAVFVGLVLLTLSPHYKMFLARLIWWLQYF  849

Qy   773  WQRWEN-----WFWNVTLRPER-FFLVLVCFPGATYDALVTFCVCHVALLCL-------818
          :   :      |:|::     | :||:|: || |::|:| ||    ||||||
Db   850  ITRAEAAHQVWIPPLNVRGGRDAILLTC---AAYPELI-FDITKILLAILGPLMVLQAG  905
```

```
Qy  2311  YVTEPRDAELRKQKVTINRQPLFPPSYHKQVRLAXEKASKVVGVNMDYDEVAAHTPSKSA  2370
          | |  :    :|: || :|  :  ||||:|||| : ||||||| || || |||:||||||
Db  2480  YCTTSKSASLRAKKVTFDRMQALDAHYDSVLKDIKLAASKVTARLLTLEEACQLTPPHSA  2539

Qy  2371  KSHITGLRGTDVR--SGAARKAVLDLQK-CVEAGEIPSHYRQTVIVPKEEVFVKTPQKPT  2427
          :| : |||| | |  :| |::|||| || |||  :|| |:| |  ||:|||  |||:||
Db  2540  RSKY-GFGAKEVRSLSGRAVNHIKSVWKDLLEDTQTPI---PTTIMAKNEVFCVDPTKGG  2595

Qy  2428  KKPPRLISYPHLEMRCVEKMYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMWS--PD  2484
          || ||:| ||:| |:|||||:  : :||||||||:|||| :  | ||: ||| |:   |
Db  2596  KKAARLIVYPDLGVRVCEKMALYDITQKLPQAVMGASYGFQYSPAQRVEFLLKAWAEKKD  2655

Qy  2485  AVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGRE  2544
             | ||| ||||| | ||: || |:| || || :  |:| :: |:||:||||:  :||
Db  2656  PMGFSYDTRCFDSTVTERDIRTEESLYRACSLPEEAHTAIHSLTERLYVGGPMFNSKGQT  2715

Qy  2545  IGYRRCRSSGYYTTSSSNSLTCWLKVNAAAEQAGMKQNPRFLICGGDCTVIWKSAGADADK  2604
          :||||||::| :|||  || |||:|: ||: : || ::| | :|||| | | |:|:: :
Db  2716  CGYRRCRASGVLTTSMGNTITCYVKALAACKAAGIIAPTMLVCGDDLVVISESQGTEEBDE  2775

Qy  2605  QAMRVFASWMKVMGAPQDCVPQKYSLEELTSCSSNVTSGITKSGKPYYFLTRDPRIPLG  2664
            :|  |: | ::   || |    ::::|:::|||||    :| | || |||||  | |
Db  2776  RNLRAFTEAMTRYSAPPGDPPREYDLELITSCSSNVSVALGPQGRRRYYLTRDPTTPIA  2835

Qy  2665  RCSAEBGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYT  2724
          |   |  | :| | :|: | | | |: |:   ::| : ::  |  :  | | | | | |:
Db  2836  RAAWETVRHSPVNSWLGNIIQYAPTIWARMVLMTHFFSILMAQDTLDQNLNFEMYGAVYS  2895

Qy  2725  VPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKR  2784
          | : ||| || :::| |||| :|:|    : |   | :   |||||: :|||| ||::
Db  2896  VSPLDLPAIIERLHGLDAFSLHTYTPHELITRVASALRKLGAPPLRAWKSRARAVRASLIS  2955

Qy  2785  RGGAHAKLARFLL-WHATSR----PLPDLDKTSVARYTFNYCDVYSPEGDVFITPQRRL  2839
          |||  :  :|: | |  |||    |:||| :||  || |:   ::   |:  :  ||
Db  2956  RGGRAAVCGRYLFNWAVKTKLKLTPLPEARLLDLSSWFT----VGAGGGDIYHSVSRAR  3010

Qy  2840  QKFLVKYLAVIVFALGL  2856
          :: | :|| ::|| | |
Db  3011  PRLLLLGLLLLFVGVGL  3027
```

RESULT 3
GNWVTW
GNWVTW
genome polyprotein - hepatitis C virus (strain Taiwan)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A40244
R;Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Liu, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
A;Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the
A;Reference number: A40244; MUID:92230206; PMID:1314449
A;Accession: A40244
A;Molecule type: genomic RNA
A;Residues: 1-3010 <CHE>
A;Cross-references: UNIPROT:P29846; GB:M84754
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
F;1-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-729/Product: envelope protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207

```
Qy  1942 ARVTQILSGLT----ITQLLRRIHQWINEDCSTPCSGSWLRDVWDWICTVLA----DFKT 1993
         :: :  :              ||  :         :::::         ::  :
Db  1890 GVQS--MVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTC 1947

Qy  1994 WLQSKLLPRLPGVPFFSCQRGYKGVWRGDGIMQTTCPCGAQLTGHVKNGSMRIW-GPKTC 2052
         :| ||        | |:|| ||||| || |||| |||||| | |||| |||   ||||
Db  1948 SNYWRGAVPVNARLCGSARPDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-C 2005

Qy  2006 FTQVP-------------------PTLRAAVA------VDGVQVCYLGEPKTPWT 2036
          |  |                  || ||         :||  ||| |||| | |
Db  2053 SNTWHGTFPINAYTTGPCTPSPAPNYSRALWRVAAEEYVEVRRVGDFHVTGMTTDNVKC 2112

Qy  2037 TSACCYGPDGKGKTVKL--PFRVDGHTPGVRM            QLNLRDALETNDCN 2080
          ::||  |    ::  :                                |: :   ||
Db  2113 PCQVPAPEFFTEVDGVRLHRYAPACKPLLREEVSFQVGLNQYYVGSQLPC----EPE---- 2165

Qy  2081 STNNTPSDEAAVSA-LVFKQEL-----------RRTNQLL------EAISAGVDTTKLPAPSIEE 2127
          :   | ||        :|:|                 || |       ::|  ||  ||    ::
Db  2166 ----PDVAVLTSMLTDPSHITAETAKRRLARGSPPSLASSSASQLS--ALSLKAAC 2215

Qy  2128 VVVRKRQFRARTGSLTLPP--PPRSVPGVSCPESLQRSDPLEGPSNLP------PSPPVL 2179
          :|  |:          ||      |:               || |:|           ::  :
Db  2216 TTRHTPPDADLIEANLLWRQEMGGNITRVESENKVVILDSFDPLRAEEDEREVSVPA--E 2273

Qy  2180 QLAMPMP-------LLGAGECNPFTAIGCAMTETGG------------GPDDLP------- 2214
          :|   | :          | |  :          |                  :||:|
Db  2274 ILRKSRKF------------------LLEPWKDPDYVPPVVHGCPLPPVK 2320

Qy  2215 ------SYPP---KKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKK 2265
                | ||      |: :| |:| |  |   ||  :|| :  :|    :       |
Db  2321 APPIPPPRRKRTVLTESTVSSALAELATKTFGSSESSAAGSGTATAPPDQPSDDGDAGS 2380

Qy  2266 KLGKSEFSCSMSYTWTD-VISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQK 2324
          :   |   |||||||   ||    | | || ::|  |                         :
Db  2381 DVESCSSMPPLEGEPGDPDLSDGSWSTVS--------------------------E 2410

Qy  2325 VTINRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVR- 2383
          :  |:   |: :       |||  |  |  ::||| |||  |    |  |    | | ::
Db  2411 EDGEGVICCSMSYTWTGALIITPCAAEESKLPINALSNSLLRHHNMVYATTSRSASQRKK 2470

Qy  2384 -SGAARKAVLDLQK-CVEAGEIPSHYRQTVIVPKEEVPVKTPQKPTKKPPRLISYPHLEM 2441
           :    | ||  |:     ||  |:   |  |  |:|:| |  ||   | :|    |
Db  2471 VTIDRLQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPHSARSKF-GYGAKDVRN 2529

Qy  2442 RCVEKMYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMWSPD--AVGATCDTVCFDST 2498
          | ||| |||: | |:| ||   || :|  |  |:   :    |    :|: | ||||||
Db  2530 LSGKAINHIRSVWKDLLEDTETPI----DTTIMAKGNEVFCVQPEKGGRKPARLIVFFPDLGV 2586

Qy  2499 ITPEDIMVETDIXYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTT 2558
          :|      |:  ::||| |:|: |||| :| || | |||| : |   ||  | |:||| ||
Db  2587 RVCEKMALYDVVSTLPQAVMGSSYGFQYSPGQRVEFLVNAWKSKKCPMGFSYDTRCFDST 2646

Qy  2559 SSSNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMG 2618
           ::| |||  | | ||  | |   ||:||||| ||:: |||  :|| |:|:||:    | |
Db  2647 VTESDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKQNCGYRRCRASGVLTT 2706

Qy  2619 APDQCVPQPKYSLEELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAA 2678
          : | | |:   ::||:|:|||||:: ||: ||||:||||||  ||| | | |:  :|
Db  2707 SCGNTLTCYLKASAACRAAKLQDCTMLVCGDDLLVICESAGTQEDAASLRVFTEAMTRYS 2766

Qy  2679 WIGYLIHHYPCLMVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVH 2738
          | | :: :||  | |: ||  : |:: :    ||  | ||     |  :     || |
Db  2767 APDGDLPQEYDQELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAWATARHTPVNS 2826

Qy  2739 GIEAFSVVRYTNAELIRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL- 2797
          | ||: :     ::|| ::| |:|: :|||||::::|| | || | | |  ||  :| |
Db  2827 WLGNIIMYAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLH 2886

Qy  2798 WHATSR-----PLP-----DLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLA 2848
          |  :||      | |     | || ||  || |  ||| |||||| ||||| | |||||||
Db  2887 GLSAFSLHSYSPGEINRVASCLRKLGVPPLRAWRHRARSVRAKLLSQGGRAATCGRYLFN 2946
```

```
Qy   819 TSSAASFFGTDSRVRAHRMLVRLGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPN 878
          :: :  ||| ||    ::   | ||         | | | |:| |  |    | |
Db   906 LTRIPYFVRAQGLIRA-CMLVR----KAAGGHYVQMALMKLAALTGTVYVYDHL 953

Qy   879 DFASKLPLQ----------EPFFPPFEGKARVYRNEGRRLACGDTVDGLPVVARLGD 924
          ::      |           | | ||    :   :: | ||| |  | || :|| | |
Db   954 -----TPLQDWAHTGLRDLAVAVEPVVFSDMETKIITWGADTAACGDIILGLPVSARRGR 1008

Qy   925 LVFAGLAMPPD------GWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSL 979
          | | | ||        |:|:|| ||  | | ||||::||:|||| || |  | ||| |||
Db  1009 EILLG---PADSLEGRGMRLLAPITAYAQQTRGLFGCIITSLTGRDKNQVEGEVQVVSTA 1065

Qy   980 ATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN-DQDI--YQPPCGAGSLTR 1036
          : :  |  |:| || |:||:|   || | |    || |  |    |    ::|  ||||||
Db  1066 TQSFLATCINGVCWTVYHGAGSKTLAGPKG---PITOMYTNVDQDLVGWHAPQGARSLTP 1122

Qy  1037 CSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPLLCSSGHVIGMFTA 1096
          | || ::||:||| | ||| :  ||  ||||| | | | ||||||||| |||| ||| ||
Db  1123 CTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPSGHVGIFRA 1182

Qy  1097 A---RNSGGSVSQIRVRPLVCAGYHPQYTAHATLDKTPTVPNEYSVQILIAPTGSGKSTK 1153
          |    :   :  : : | || | :  ::  :    :  ||:: | :| ||||||||||||
Db  1183 AVCTRGVAKAVDFVPVBESMETMRSPVFTDNS---SPPAVPQAFQVAHLHAPTGSGKSTK 1239

Qy  1154 LPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGMY 1213
          :|  | |:||:|||||||| |: |   |  || ||:|| :   | |  ||:||||||| :
Db  1240 VPAAYAAQGYKVLVLNPSVAATLGFGAYMSKAHGVDPNIRTGVRTIITTGAPITYSTYGKF 1299

Qy  1214 LT-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPH 1271
          |   | | |  ||:| ||||:||:||:|||| || :|   | ||||||||||| :  ||
Db  1300 LADGGGCSGGAYDIIMCDECHSTDSTTILGIGTVLDQAETAGARLVVLATATPPGSVTVPH 1359

Qy  1272 ANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKGHCDELANELARKGITAVSYYR 1331
          :|| || | :|  ||:||  |  |:  ||||||| |:||: ||||| :|: |||:|| ||||
Db  1360 PNIEEIALSNTGEIPFYGKAIPIETIKGGRHLIFCHSKKKCDELAAKLSALGIHAVAYYR 1419

Qy  1332 GCDISKIP-EGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGV 1390
          | |:| ||  | |||||||||  |:|||||||:|| : ::||||||||||||: ::   :
Db  1420 GLDVSVIPASGNVVVVATDALMTGFTGDFDSVIDCNTCVTQTVFDFSLDPTFTIETTMPQ 1479

Qy  1391 SAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTIL 1450
           |:  :|||||||||||||| ||  || |||| || :|||||||| ||| : |:   ||
Db  1480 DAVSRSQRRGRTGRGRAGIYRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAETSVRL 1539

Qy  1451 DTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLCHQYGY 1509
          : |  ||||| : |: : || || | :      || | :|  || |  ||:   |:
Db  1540 RAYLNTPGLPVCQDHLEFWESVPTGLTHIDAHFLSQTKQAGDNFPYLVAYQATVCARAQA 1599

Qy  1510 AAPNDAPRWQGARLGKK-----PCGVLWRLDGADACPGPEPSEV------TRYQM-CFT- 1556
          : | |  | |  ||:       |||:|| || |  ||||| | :       ::   |
Db  1600 PPPSWDQMWKCLTRLKPTLHGPTPLLYRL------GAVQNEVTLTHPITKYIMACMSAD 1652

Qy  1557 -EVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEEI---- 1611
            |:|: | | | || |  :||   :  | | | |:      | |:|:| ||
Db  1653 LEVVTS-TWVLVGGVLAALAAYCLTTGSVVIVGR------IIILSGKPAVVVDREVLYQEF 1705

Qy  1612 --VEECASFIP-LEAMVAAIDKLK------STIT---TTSPFTLETALEKLNTFLGPH 1657
            :||||| |  :| |: ||   |        |:    | | || |:: | ||  :
Db  1706 DEMEECASHLPYIEQGMQLAEQFKQKALGLLQQTATKQAEAAAP-VVESKWRTLEAFWAND 1764

Qy  1658 AATILAIIEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDAR 1717
          ::   |  | | || ||| |||| |:  |  | | :| : :  :|  ||| ||:  |
Db  1765 MWNFISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTTQSTLLFNILGGWAAQLAPPG 1824

Qy  1718 GALAFMWAGAAGTALAGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAG 1774
          :| ||: |:|:|:| || :  : |    ::|::||  || ||| || :::|:  |::
Db  1825 AASAFVGAGIAGAAVG---SIGLGKVLVDMVAGYGAGVAGALVAFKVMSGEMPSTEDIVN 1881

Qy  1775 LVYSAFNPAAGVVGVGVGVVCSACAMFALTTAGPD-HWPNRLLTMLARSNTVCNEYFIATRDIR 1832
          | ::|:  :||:|||:|: |  :|| | | | ||    ||||| :|| | || |  |: |:| |
Db  1882 LLPAILSPGALVVGVVCAAILRRHVDPEGAVOWMNRLLIAFASRGNHVSPTHVYPESDAA 1941

Qy  1833 RKILGILEASTPWSVISACIRWLHTPTEDDCGLI---AWGLEIWQYVCNFFVICFNVLKA 1889
```

```
Db     2947  WAVKTKLKLTPIPAASQLDLSKWFVAGY----------GGGDIYHSLSRARPRW------ 2990
                                                    TSGNPIRVPTGC  358
Qy     2849  VIVFALGLIAVGLAI 2863
                   ||:::|||:||
Db           :
Qy     2991  -FMLCLLLLSVGVGI 3004

RESULT 4
A45573
genome polyprotein - hepatitis C virus (strain JT)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: A45573
R;Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata,
Virus Res. 23, 39-53, 1992
A;Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: a
A;Reference number: A45573; MUID:92295714; PMID:1318627
A;Accession: A45573
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-3010 <TAN>
A;Cross-references: UNIPROT:Q00269; GB:D11168; GB:D01171; NID:g2221612; PIDN:BAA01943.1;
A;Experimental source: HCV-JT
A;Note: sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;710-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match       21.0%;   Score 3232;  DB 1;  Length 3010;
Best Local Similarity  29.9%;  Pred. No. 1.3e-192;
Matches 971; Conservative 451; Mismatches 1187; Indels 640; Gaps 115;

Qy       8  TSPVPAPRTRKN---------VSIKTSVERGQRAKRKVQRDARPRNYK  54
               |:| :|  :|:      |::| | |  :|  |:: |||| ||
Db       3  TNPKPQRKTKRNTYRRPQDVKFPGGGQIVGGVYVLPRRGPTLGVRATRKTSERSQPRGRR  62

Qy      55  IAGIHDGLQTLAQAALP-----------------------------AHGWGRQDPRH  82
               |||| |  ||:|                                 |||| |||
Db      63  ---------QPIPKARRPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPTDPRR  114

Qy      83  KSRNLGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG---WFG  139
              ||||| ||||| |:|  :||: ||||  || |: |: |||||||||| |||    |||
Db     115  RSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLPGCS  173

Qy     140  VHLFVVCLLS-LACPCSGARVTDPDTNTTILTNCCORNQVIYCSPSTCLHEPGCVIC---  195
              |:|  || :  |||   | :| |  |||:||: ::||||:|:||| |||| ::
Db     174  FSIFLLALLSCLTIPASAYQVRNA-SGLYHVTNDCSNSSIVYEAAGMIMHTPGCVPCVRE  232

Qy     196  --ADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWL  253
                ||:||  | |:| ||  || |:| ::  ||||| ||| |||   |::|||| ||| |
Db     233  NNASRCWVALTPTLA-ARNTSIPTTTIRRHVDLLVGAAAFCSAMYVGDLCGSVFLVSQLF  291

Qy     254  V---RHWLIHIDLNETGTCYLEVPTGIDPGFLG--FIGW---MAGKVEAVIFLTKLASQV  305
              |    ::|:|| | |:| |: ||||  |  ||     |||   ||| || ||| |:|||
Db     292  TFSPRRYETVQDCN---------CSIYPGHVSGHRMAWDMMNWSPTTALVVSQLL-RI  340

Qy     306  PYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEA-----------------  346
              | | |:|||| |:||| |: |::||| |: || |:|::||
Db     341  PQAVVDMVAGAHWGVLAGLAYYSMVGNWAKVLIVMLLFAGVDGVTYTTGGSQARHTQSVT  400
```

```
Db      347  ------------------------------------------------------
Qy      401  SFFTQGPAQRIQLINTNGSWHINRTALNCNESLNTGFFAALFYAHKFNSSGCPERMASCS  460

Qy      359  SIABFC---SPLMIPCP----------------CHSY----LSENVSEVICYSPKWTRPTTLE  398
                :| ||    ||:|| |                    |:::
Db      461  SIDKFAQGWGPITYTEPRDLDQRPYCWHYAPRQCGIVPASQVCGPVVCFTPS---PVVVG  517

Qy      399  YNN-----SISWYPYTIPGARGCMVKFXNNT------WGCCIRRNVPSY------CTM  439
              :       : ::  |::|   |::::|||        ||   :
Db      518  TTDRSGAPTYNW------GANETDVLLLNNTRPPQGNWFGCTWMNSTGFTKTCGGPPCNI  571

Qy      440  G--------TDAVNDTRNTYEACGVTPWLT-----------------TAWHNGSALKLAILQ---  477
                         |   |   || ||| | |||                :
Db      572  GGVGNLTLTCPTDCFRKHPEATYTKCGSGPWLTPRCIVDYPYRLWHYPCTVNFTIFKVRM  631

Qy      478  YPGSKE--MFKPHNWMSG-HLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGS  534
              ||| |:    | ||::   || ||||| :| |:|   ||| |  :| |  |:|||:
Db      632  YVGGVEHRLSAACNWTRGERCDLEDRDR-------SELSPLLLSTTEWQTLPCS-------  678

Qy      535  WLQVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLGSK  594
              ||:|       ::|   |:|:   :|  | :|:   || | ||:
Db      679  ----------FTTLPALSTGLIHLHQNIVDVQVLYGIGSAVVSFVIKWEYIVLFLLLADAR  730

Qy      595  YLILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYLRCR  654
                                                          :: |::|:
Db      731  VCACLWMMLL---------------------------------------IAQEAALENLVV------  753

Qy      655  LRYAALLGFVPMAAGLPLTFVAAAAAQPDYDWVR-LLVAGLVLWAGRNRGHRIALLVG  713
               || | ||| : |:|  | | |         :     ||:| ||| | : | ::|||
Db      754  LNAASLAG----ADGI-LSFLVFFCAA-------WYIKGRLVPGAAY------ALYG  792

Qy      714  PWPLVALLTLHLVTPASAFDTEI---IGGLTIPPVVALVVMSRFGFFAHLLPRCALVNS  770
              ||||  ||| |:| || |::     |||| |||||| |:::| |  ||| ||| |||
Db      793  VWPL--LLLLLALPPRAYAMDREMAASCGGVVFVGLILLTLSPHYKVF---LARLIWWLQ  847

Qy      771  YLWQRWEN----WFWNVTLRPERFFLVLV---CFPGATYDALVTFCVCHVALLCLTSSAA  823
                ||| ||    || ||  || | |:| :    : |||  |  |:  |  | ||: ||
Db      848  YFITRAEAHLCVWPPLNVRGGRDAIILLTCAAHPELIPDITKLLLAILGPLMVLQAAIT  907

Qy      824  S---PFGTDSRVRAHRMLVRLGKCHAWYSHYULKFFLLVFGENGVFFYKCHLHGDVLPNDF  880
                   ||  :   |:|| | :|       :|:||  |||| ||    ||| || | |
Db      908  AMPYFVRAQGLIRA-CMLVR----KVAGGHYVQMAFMKLAALTGTYVYDHL----------  953

Qy      881  ASKLPLQ------------------EPFPFPEGKARVYRNEGRRLACGDTVDGLPVVARLGDLV  926
                 |||:                   || | ||:  ::| : |: |||     : |
Db      954  ---TPLQDWAHAGLRDLAVAVEPVVFSDMETKIITWGADTAACGDIILGLPVSARRGREI  1010

Qy      927  FAGLA--MPPDGWAITTAPPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSYM  984
              |||    :||  |  |  |:|  |  ||| |||| |::|| |  ||     | |
Db     1011  LLGPADSIEGQGWRLLAPITAYAQQTRGLLGCIVTSLTGRDKNQVEGGVQVVSTATQSFL  1070

Qy      985  GFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN-DQDI--YQPPCGAGSLTRCSCGE  1041
              ||  ||:|   :|| :|:  |::|   :|::|  :  :|  | |: | | :|: :| | |
Db     1071  ATCVNGVCWTVFHGAGSKTLAGPKG---PITQMYTNVDQDLVGWHAPPGARSLTPCTCGS  1127

Qy     1042  TKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAA---R  1098
               | |||| ||| :|  :|   ||| || | || |||| ||| ||:| :|    :|
Db     1128  SDLYLVTRHADVIPVRRGDGRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTR  1187

Qy     1099  NSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQLLIAPTGSGKSTKLPLSY  1158
              | ::| |:||:  ||| |  || | | || ||| ::|  :|:||   || || : | |
Db     1188  GVAKAVDFIPVESMETMRSPVFTDNS---SPPAVPQTFQVAHLHAPTGSGKSTKVPAAY  1244

Qy     1159  MQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYLT-GA  1217
              |:|| ||:|||| |||  :||:| :  || ||||: |: |   || |:| || |
Db     1245  AAQGYKVLVLNPSVAATLGFGAYMSKAHGTDPNIRTGVRTITTGAPITYSTYGKFLADGG  1304

Qy     1218  CSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATEATPPGVIPTPHANITE  1276
              |||  ||:|:|||| :|:: ::||||::|   | || |:||||| |||| |  :|::
Db     1305  CSGGAYDIIICDECHSTDSTTILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEE  1364

Qy     1277  IQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELARKGITAVSYYRGCDIS  1336
```

```
Db  1365  VALSNTGEIPFYGKAIPLEAIKGGRHLIFCHSKKKCDELAAKLSGLGINAVAYYRGLDVS  1424
Qy  1337  KIP-EGDCVVVATDALCTGYTGDPDSVYDCSLMVEGTCHVDLLDPTFTMGVRVCGVSAIVK  1395
Db  1425  VIPTSGDVVIVATDALMTGYTGDFDSVIDCNTCVTQTVFSLDPTFTIETTVPQDAVSR    1484
Qy  1396  GQRRGRTGRGRAGIYYVVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTILDTYRT   1455
Db  1485  SQRRGRTGRGRGGIYRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYLN   1544
Qy  1456  QPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLCHQYGYAAPND    1514
Db  1545  TPGLPVCQDHLEFWESVPTGLTHIDAHFLSQTKQAGDNFPYIVAYQATVCARAQAPPPSW   1604
Qy  1515  APRWQG-ARLGKK--PCGVLWRLDGADACPGPEPSE-----VTRYQM-CFT---EVNT    1560
Db  1605  DQMWKCLIRLKPTLHGPTPLLYRL------GAVQNEITLTHPITTKFIMACMSADLEVVT   1657
Qy  1561  SGTAALAVGVGVAMAYLAIDFGATCVRRCWSITSVPTGATVAPVVDEEEI------VEE   1614
Db  1658  S-TWVLVGGVLAALAAYCLTTGSVIVGR------IILLSGRPAVVPDREVLYREFDEMEE   1710
Qy  1615  CASFIP-LEAMVAAIDKLK-------STIT----TTSPFTLETALEKLNTFLGPHAATIL  1662
Db  1711  CASHLPYIEQGMQLAEQFKQALGLLQTATKQAEAAAP-VVESRWRALEAFWAKHMWNFI    1769
Qy  1663  AIIEYCCGLVTLPDNPFASCVPAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAF  1722
Db  1770  SGIQYLAGLSLTPGNPAIASLMAFTASTTSPLTTQNTLLFNILGGWVAAQLAPPSAASAF  1829
Qy  1723  MMAGAAGTALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSA  1779
Db  1830  VGAGIAGAAIG--SIGLGKVLVDILACYGAGVAGALLVAFKVMSGEAPSAEDIVNLLPAI  1886
Qy  1780  FNPAAGVGVLSACAMFALTTAGPD--HWPNRLLTMLARSNTVCNEYFIATRDIRRKI     1835
Db  1887  LSPGALVVGVV--CAAILRRHVGPGEGAVQWNRLIAFASRGNHVSPTHYVPESDAAARV   1944
Qy  1836  LGILEASTPWSVISACIRWLHTPTEDDCGLI--AWGLEIWQVCNFPVICFNVLKAGVQ    1892
Db  1945  TQILSSLT----ITQLLKRLHQWINEDCSTPCSGSWLKDVWDWICT---VLTDFKTWLQ   1996
Qy  1893  S--MVNIPGCPFYSCQKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNY     1950
Db  1997  SKLLPKLPGVPFFSCQRGYKGVWRGDGIMQTTCPCGAQITGHVKNGSMRIV-GPKTCSNT  2055
Qy  1951  WRGAVPVNARLCGSARDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFTQ   2008
Db  2056  WHGTFPINAYTTGPCTPSPAPNYSRALWRVAAEEYVEITRVGDFHYVTGMTTDNVKCPCQ  2115
Qy  2009  VP-------------------------PTLRAAVAVD-------GVQVQCVLGEPKTPWTTSA  2039
Db  2116  VPAPEFFTELDGVRLHRYAPACRPLLREDVTFQVGLNQYLVGSQLPC---EPE------   2165
Qy  2040  CCYGPDGKGKTVKL-PFRVDGHTPGVRMQLNLRDALETND---------CNSTNNT     2085
Db  2166  ----PDVAVLITSMLTDPSHITAETAKRRLARGSPPSLASSSASQLSAPSLKATCTTHHDS 2221
Qy  2086  PSDEAAVSALVFKQEL-----RRTNQLL-----EAISAGVDTTKLPAPSIEEVVRKR    2133
Db  2222  PDADLIEANLLWRQEMQGNITRVESENKVVILDSFDPLRAEEDEREV---SVAAEILRK-  2277
Qy  2134  QFRARTGSLTLPP---PPRSVPGVSCP--ESLQRSD---------PLEGPSNLPPSPP--- 2177
Db  2278  ---SKKFPPALPIWARPDYNPPLLESWKSPDVYPPAVHGCPLP-PTTGPPIPPPRK     2329
Qy  2178  ----VLQLAMPMPLLGAGECNPFTAIGCAMTETG---GGPD------DLPSY---P     2217
Db  2330  KRTVVLTESTVSSALAELATKTFGSSGSSAVDSGTATAPPDQTSDDGDKESDVESYSSMP  2389
Qy  2218  P---KKEVSEWSDEWSWSTATTASSYVTGPPYKIRGDKTQSAPARKPTKKKLGKSEFSC   2274
```

```
Db  2390  PLEGEPGPDDLSDGSWSTVSGEAS-------------------DDIVCC         2419
Qy  2275  SMSYTWTD-VISFKTASKVLSATRAITSGFLKQRSLVVTEPRDAELRKQKVTINRQPLF   2333
Db  2420  SMSYTWTGALITPCAAEESKLPINALSNSLLRRHNMVYATTSRSASLRQKKVTFDRLQVL  2479
Qy  2334  PPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLD  2393
Db  2480  DDHYRDVLKEMKASTVKAKLLSVEEACKLTPPHSAKSKF-GYGAKDVRN-LSSKAINH    2537
Qy  2394  LQK-----CVEAGIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMY   2449
Db  2538  IRSVWKDLLEDTETPI---DTTIMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMAL  2594
Qy  2450  GQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMW----SPDAVGATCDTVCFDSTITPEDI  2504
Db  2595  YDVVSTLPQAVMGSSYGFQYSPGQRVEFLVNAWKSKKSP--MGFSYDTRCFDSTVTESDI  2652
Qy  2505  MVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSL  2564
Db  2653  RVEESIYQCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTL  2712
Qy  2565  TCWLKVNAAEAEAQGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCV 2624
Db  2713  TCYIKATAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDP  2772
Qy  2625  PQPKYSLELTCSSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLI   2684
Db  2773  PQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNII  2832
Qy  2685  HHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNVTVPVEDLPSIIAGVHGIEAFS  2744
Db  2833  MYAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGLSAFS  2892
Qy  2745  VVRYTNABILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WHATSR  2803
Db  2893  LHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCKYLFNWAVRTK   2952
Qy  2804  ----PLP---DLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFAL    2854
Db  2953  LKLTPIPAASQLDLSSWFVAGYS--------GGDIYHSLSRARPRW------FMWCL    2995
Qy  2855  GLIAVGLAI  2863
Db  2996  LLLSVGVGI  3004
```

RESULT 5

S40770

genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructur
    protein NS4a; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: S40770; PC1285
R;Okamoto, H.
submitted to the EMBL Data Library, March 1992
A;Reference number: S40770
A;Accession: S40770
A;Molecule type: genomic RNA
A;Residues: 1-3011 <OXA>
A;Cross-references: UNIPROT:Q03463; EMBL:D10749; NID:g221586; PIDN:BAA01582.1; PID:g2215t
R;Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda,
Jpn. J. Exp. Med. 60, 167-177, 1990
A;Title: The 5'-terminal sequence of the hepatitis C virus genome.
A;Reference number: PC1284; MUID:91013116; PMID:2170712
A;Accession: PC1285
A;Molecule type: genomic RNA
A;Residues: 1-513 <OK2>
A;Cross-references: GB:D00831; NID:g221511; PIDN:BAA00705.1; PID:g221512
A;Experimental source: isolate HC-J1
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine

F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1119/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match   Length 3011;
Best Local Similarity   Score 3224.5; DB 1; Pred. No. 3.8e-192;
Matches '972; Conservative 434; Mismatches 1232; Indels 583; Gaps 112;

```
Qy   660  LLGFVPMAAGLP--LTFFVAAAAQPDYDMWVRLLVAGLVLWAGRNRGHRIALLVGPWPL  717
               ||||       ||:|||  ::|  |                  ||||||||
Db   759  LAG---TRGLVSFLVFFCFA------WYLK---------GRWVPGAAYALYGWMPL  796

Qy   718  VALLTLHLVTPASAFDTEI---IGGLTIPPVALVVMSRFGFFAHLLPRCALVNSYLWQ  774
               ||| |  ||| ||||  :|:|||:||| : ||:|
Db   797  ---LLLLLALPQRAYALDTEVAASCGGVVLVGLMALTLSPYY------KRC--ISWCLW-  844

Qy   775  RWENWFWNVTLRPERFFLVLVCFPPGATYDALVTFCVCHVALLC----------LTSSA  822
               |||:||||||  :|:| :|  :|  ::    :  |  |  |          : ::
Db   845  -WLQYFLTRVEAQLHVWVPPLNVRGGRDAVILLMCVVHPTLVFDITKLLLAVLGPLWILQ  903

Qy   823  ASFFGTDSRVRAHRMLVRLGKCHAWY-SHYVLKFFLLVFGENGVFFYKHLH--GDVLPND  879
               || :| |  ::    ||||||  ||:||::  |||| :||:||  |||:  ||||||
Db   904  ASLLKVPYFVRVQGLLRICALARKMVGGHYVQMAIIKLGALTGYVVNHLTPLRDWAHNG  963

Qy   880  FASKLPLQEPFFPFEGKARVYRNEGRRLACGDTVDGLPVVARLGDLVFAGLA--MPPDGW  937
               ::|  ||||   |  |||  :| || |       : |:  :|  :|| |     |||||
Db   964  LRDLAVAVEPVVFSQMETKLLTWGADTAACGDIINGLPVSARKGREILLGPADGMVSKGW  1023

Qy   938  AITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIPRLGSLATSYMGFVCDNVLYTAHH  997
               : ||||    |||| | ||   |     |||  |  |: |||||||||:  :  || ||
Db   1024 RLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYH  1083

Qy   998  GSGKRRLAHPTGSIHPITVDAANDQDI--YQPPCGAGSLTRCSCGETKGYLVTRLGSLVE  1055
               |:||||:||||||||||||||:|:|      :|||||||||||   |    ::  ||
Db   1084 GAGTRTIASPKGPV--IQMYTNVDQDLVGWPAPQGARSLTPCTCGSSDLYLVTRHADVIP  1141

Qy   1056 VNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAA---RNSGGGSVSQIRVRPL  1112
               |     |||||||    ||||||| || :  || :|          :|||    : |  |
Db   1142 VRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHVGIFRAAVCTRGVAKAVDFIPVESL  1201

Qy   1113 VCAGYHPQYTAHATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSV  1172
               | ||: |  ||  ||||||::|  |:  |||||||||||| ||  |:::||:||||||||
Db   1202 ETTMRSPVFTDNS---SPPAVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSV  1258

Qy   1173 ATTASMPKYMHATGVNPNCYFNGKCTNTGASLTYSTYGMYLT-GACSRN-YDVIICDEC  1230
               | |:|| :|| |||||||||:| ||||:||||:||||||  : ||| ||  |:|||||||
Db   1259 AATLGFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCGSGAYDIIICDEC  1318

Qy   1231 HATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDEGTIPFHGK  1290
               ||||||||||||||||||||||||||||||||||||||:|||||:||||| || ||:|||
Db   1319 HSTDATSVLGIGTVLDQAETAGARLVVLATATPPGSITVPHANIEEVALSTTGEIPFYGK  1378

Qy   1291 KIKEENLKKGRHLIFEATKKGHCDELANELARKGITAVSYYRGCDISKIP-EGDCVVVATD  1349
               :| | |  || ||||  |  |   :| |||  |::||:|||| | :|    |||||||:||
Db   1379 AIPLEAIKGGRHLIFCHSKKKCCDELAAKLVALGVNAVAYYRGLDVSVIPTSGDVVVVATD  1438

Qy   1350 ALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGI  1409
               ||  ||||||||| ||  :||| |  ||||  | ||| ||||: | | |||||:||||||
Db   1439 ALMTGYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTLPQDAVSRTORGRGRGKPGI  1498

Qy   1410 YYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEW  1469
               | | || | || |  || || |:| ||| |   :|    ||  |  | ||  |   |  |
Db   1499 YRFVAPGERPSGMFDSSILCECYDTGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFW  1558

Qy   1470 ADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLCHQYGYAAPNDAPRWQG-ARLGKK-  1526
               |  :   |:   |||| :|  |||:|||  |:|  ||||  |  |  || :  || |
Db   1559 EGVFTGLTHIDAHFLSQTKQGGENFPYLVAYQATVCARAQAPPSWDQMWKCLIRLKPTL  1618

Qy   1527 --PCGVLWRLDGADACPGPEPSEVTRYQM-CFT---EVNTSGTAALAVGVGVAMAYLAID  1580
               || ||  | | ||| || | |   :|       |||||| |  | | || |
Db   1619 HGPTPLLYRL-GAVQGEVTLTHPVTKYIMTCMSADLEVVTS-TWVLVGGVLAALAAYCLS  1676

Qy   1581 TFGATCVRRCWSITSVPTGATVAPVVDEEI------VEECASFIPL-EAMVAA---ID  1629
               |||  ||      :|  || |   |: :||         ||  |:| :|: |
Db   1677 TGCVVIVGR------IVLSGRPAIIPDREVLXREFDEMEECSQHLPYIEQGMMLAEQFKQ  1730

Qy   1630 KLKSTITTTS-------PFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFASC  1682
               | :::                  :  | |||:   :  |  |: ::| |:|:|||:| |
Db   1731 KALGLLQTASRQAEVIAP-TVQTNWQKLEAFWAKCHMWNPISGTQYLAGLSTLPGNPAIAS  1789

Qy   1683 VFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF--  1740
               ::
Db   1683 ---WMMLL----------ISQAEAALENLVI------LNAAS  758
```

Db   2797 TGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMFAPTLWARMLLMTHFFSVLIAR 2856

Db   2708 DKLPETVTFPDWYGKNYTVPVEDLPSIIAGVGIEAFSVVRYTNAEILRVSQSLTDMTMPP 2767

Qy   2857 DQLEQALDCEIYGACYSIEPLDLPPIIQRLHGLSAFSLHSYSPGEINRVAACLRKLGVPP 2916

Qy   2768 LRAWRKKARAVLASAKRGGAHAKLARFLL-WHATSR-----PLPDLDKTSVARYTTFNVC 2822

Db   2917 LRAWRHRARSVRARLLSRGGRAAICGKYLFNWAVRTKLKLTPIAAAGRLDLSGWFTAGYS 2976

Qy   2823 DVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAI 2863

Db   2977 G-----GDIYHSVSHARPRW-------FWFCLLLLAAGVGI 3005

RESULT 6
GNMVCJ
genome polyprotein - hepatitis C virus (strain J)
N;Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural
  protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: A39253; PS0086
R;Kato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimoto
  Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990
A;Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients v
A;Reference number: A39253; MUID:91088550; PMID:2175903
A;Accession: A39253
A;Molecule type: genomic RNA
A;Residues: 1-3010 <KAT>
A;Cross-references: UNIPROT:P26662; GB:D90208; NID:g221610; PIDN:BAA14233.1; PID:g221611
R;Kato, N.; Ohkoshi, S.; Shimotohno, K.
  Proc. Jpn. Acad. 65B, 219-223, 1989
A;Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence vari
A;Reference number: PS0085
A;Accession: PS0086
A;Molecule type: genomic RNA
A;Residues: 2650-2707 <KA2>
A;Experimental source: Japanese isolate
C;Comment: The cleavage sites of this polyprotein have not been determined.
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <MEE>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,2;

Query Match   21.0%; Score 3223; DB 1; Length 3010;
Best Local Similarity   30.2%; Pred. No. 4.8e-192;
Matches 977; Conservative 439; Mismatches 1206; Indels 614; Gaps 117;

Qy   8 TSPVPAPRTRKNKQTQASYPVSIK------------------TSVERGQRAKRKVQRDARPR 51

Db   3 TNPKPQRKTKRNTNRR----PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTISERSQPR 59

Qy   52 NYK----IAGIHDGLQTLAQAALP------------------------AHGWGRQDPRHKSRN 86

Db   60 GRRQPIPKARRPEG-RTWAQPGYPWPLYGNEGMGWAGWLLSPRGSRPSWGPTDPRRRSRN 118

Qy   87 LGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG----WFGVHLF 143

Db   119 LGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLPGCSFSIF 177

Qy   144 VVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVICADE---- 198

```
Db   178  ::|||||:::||:|||||||||:: :::|| | | :::||| ||||||
          LLALLSCLTIPASAYEVRNV-SGIYHVTNDCSNSSIVYEAADMIMHTPGCVPCVRESNFS  236

Qy   199  -CWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV---  254
          ||| ::: :: :|:| ||:|  || ::||: :||:|||: ::||:|
Db   237  RCWVALTPTLAARNSSIPTTT-IRRHVDLLVGAAALCSAMYVGDLCGSVFLVSQLFTFSP  295

Qy   255  RHWLIHIDLNETGTCYLEVPTGIDPGFLG--MAGKVBAVIFLTKLLASQVPYAI  309
          :|||| : || |:||  | :||||| | |  :|||  |:|:|| |:| : :
Db        -CSIYPGHVSGHRMAWDMDMMNWSPTTALVVSQLL-RIPQAV  344

Qy   296  RRYETVQDCN-------  
          :: |   ||
Db   310  ATMFSSVHYLAVGALIYYASRGKWYQLLLALMLY-  343

Qy   345  VDMVRAGAHWGVLAGLAYYSMVGNWAKVLIVMLLFAGVDGHTHVTGRVASSTQSLVSWLS  404
          ::|                                          :||:
Db   344  -----------------IEAT---------------SGNPIRVPTGCSIAE  362

Qy   405  QGPSQKIQLVNTNGSWHINRTALNCNDSLQTGFIRALFYAHRFNASGCPERMASCRPIDE  464
          | :||: :: | :|||  |:| :|:|||:  ::|:||  |:::|   |  |:
Db   363  FC---SPLMIPCP-----------CHSY-------LSENVSEVICYSPKWTRPITILEYNN-  401

Qy   465  FAQGWGPITHDMPESSDQRPYCWHYAPRPCGIVPASQVCGPVYCFTPS---PVVVGTTDR  521
          ||  || |  | :| |  ||    | |   |  ||| | :
Db   402  --SISWYPYTIPGARGCMVKFKQNNT-----------WGCCRIRNVPSY--------CTWG---  440

Qy   522  FGAPTYSW------GENETDVLLLSNTRPPQGNWFGCTWMNSTGFTKTCCGPPCNIGGVG  575
          ||          |:|:|||:|:|  ||:: ||| |  |::||  |||||| | ||| |
Db   441  -----TDAVWNDTRNTYEACGVTPWLT-----------TAWHNGSALKLAILQ--YPGS  481

Qy   576  NNTLVCPTDCFRKGHPEATYTKCCGSGPWLTPRCMVDYPYRLWHYPCTVNFTVFKVRMYVGG  635
          |:|| |  |  :||||| | |:||||||||| |  | |  :|:||| |:  :|:  :| ||
Db   482  KE---MFKPHNWMSG-HLYFEGSDTPIVFYFDPVNSTLLPPERWARLPGTPPVVRGSWLQV  538

Qy   636  VEHRLNAACNWTRGERCDLEDRDR------SELSPLLLSTTEWQILPCS------  678
          |:  |  |     |||| |  |:   |  |||:||| | :|||  |
Db   539  PQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLIL  598

Qy   679  -----FTTLPALSTGLITHLHRNIVDVQYLYGIGSAVVSFAIKWEYILLLFLLLADARVCAC  734
          | ||| ||:|    :|: ||:   : :| |:: ||||: ::::|| :
Db   599  AYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYA  658

Qy   735  LWMMLL----------IAQABATLE-NLVVLNAASVA------GAH  763
          |  :|               :|  :|| ::|:  |:       | |
Db   659  ALLGFVPMAAGLPLTFVAAAAAQPPYDWVR--LLVAGLVLWAGRNRGHRIALLVGPWPL  717

Qy   764  GLLSF-------LVFFCAA--------WYIKGRLVPGAAY-  796
          |:| |       ||: |               |  |  |
Db   718  VALLTLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQRWE  777

Qy   797  -LLLLALPRAYAMDREMAASCCGAVFVGLVLLTLSPYYKVFLARLIIWLQYFITRAE  854
          | |||: |  |  :    | |||||:||||| |||| :|:|| | |||| :::||
Db   778  N----WFWNVTLRPERPFLVLV---CFPGATYDALVTFCVCHVALLCLTSSAAS---FFG  827

Qy   855  AHLQWVPPLNVRGGRDAIILLTCAVHPBELIFDITKLLLALIGPLMVLQAGITRVPYFVR  914
          |  | ||| |:| |: ||:|  |   |   :| :| |:| :| : |::|| |:|| |
Db   828  TDSRVRAHRMLVRLGKCHAWTSHYVLKFFLLVFGENGVFFYKHL-------HGDVLPNDF  880

Qy   915  AQGLIRA-CMLVR----KVAGGHYVQMAFMKLAALTGTYVVDHLTPLRDWAHAGL--RDL  967
          ||  ||  |           | |::  |:|  :|::| |:|:   |::|    |: | |
Db   881  ASKLPLQEPFFPFEGKARVVRNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPD---  935

Qy   936  GWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTA  995
          | | ||   ||   ||  ||:  | | ||  | | |::|| | |::  |:|  || :
Db        ----EPVVVFSDMETKLLITWGADTAACGDIISGLTGRDKNQVDGEVQVLSTATQSFLATCVNGCWTV  1021

Qy   996  HHGSKGRRLAHPTGSIHPITVDAAN-DQDI--YQPPCGAGSLTRCSCGETKGYLVTRLGS  1052
          |:|:||| :| |  ||  | | ||  |||:   |||| |  |:|||| |  | || |:
Db   1022 GWRLLAPITAYSQQTRGLLGCIITSLTGRDNQVDGEVQVLSTATQSFLATCVNGCWTV  1081

Qy   1082 YHGAGSKTLAGPKG---PITOMYTNVDQDLVGWPAPPGCARSMTPCTCGSSDLYLVTRHAD  1138
          |:| |||| ||:|    | :| | |:|| | | | | | |  | | |:||||||||:|
Db   1082 ----------------------  

Qy   1053 LVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAA---RNSGGSVSQIRV  1109
          |:| | |  |  | | | ||:|  | | |||:|  |:::|: ||    || ||:|::| :
Db   1173 SMLTDPSHITAETAKRRLARGSPPSLASSSAQSLAPSLKATCTTHHDSPDADLIEANLL  2232
```

```
Db   1139 VVPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPSGHVVGIFRAAVCTRGVAKAVDFIPV  1198

Qy   1110 RPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQEKYEVLVLN  1169
          ||:| :||||||| |:|:|| |||||: ||:|| |||||||| ::|:::| ::|||:||
Db   1199 ESMETTMRSPVFTDNS---SPPAVPQTFQVAHLHAPTGSGKSTKVPAAYAAQCYKVLVLN  1255

Qy   1170 PSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYLT-GACSRN-YDVIIC  1227
          ||:|| |:  :: |:||:| ||   || || || |||||||| ::  ::|  : ||:||
Db   1256 PSVAATLGFGAYMSKAHGIEPNIRTGVRTIITGGPITYSTYCKFLADGGCSSGAYDIIIC  1315

Qy   1228 DECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANTEIQLTDEGTIPF  1287
          |||::|:||||:|||:| | | |||  :||||| ||| |:  | ||:  : |:||:|||
Db   1316 DECHSTDSTTILGIGTVLDQAETAGARLVVLATATPPGSITVPHPNIEEVALSNTGEIPF  1375

Qy   1288 HGKKIKEENLKKGRHLJPEATKKHCDELANELARKGITAVSYRGCDISKIP-EGDCVVV  1346
          :| :| ||:| ||||| | | |:|  || ||| :|:| |||||| ||  |  :|| ||:
Db   1376 YGKAIPIEAIKGGRHLIFCHSKKKCDELAAKLTGLGLNAVAYYRGLDVSVIPTSGDVVVV  1435

Qy   1347 ATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGR  1406
          ||||: ||||||||||:|| |:::|:|:|| ||:| ::  ::|  |:| | :||||||||
Db   1436 ATDALMGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTLPQDAVSRAQRGRTGRGR  1495

Qy   1407 AGIYYYVDGSCTPSGMVPECNIVEAFDAAKAMYGLSSTEAQTILDTYRTQPGLPAIGANL  1466
          :||   ||| :   | ||   | | |:||  ||| | ||:|:| :| |:  |:|| : :|
Db   1496 SGIYRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAETSVRLRAYLNTPGLPVCQDHL  1555

Qy   1467 DEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLCHQYGYAAPNDAPRWQG--ARLG  1524
          || : ||  |:| ||| :   |:  | |::    :  :|||  | |  |  | |   | |
Db   1556 EFWESVFTGLTHIDAHFLSQTKQAGDNLPYLVAYQATVCARAQAPPSWDQWMKCLIRLK  1615

Qy   1525 KK---PCGVLWRLDGADACPGPEPSEV------TRYQM-CFT---EVNTSGTAALAVGVG  1571
          |:   |  |   :|   ||  | :|       | |:     :  :|  | |  | :| |
Db   1616 PTLHGPTPLLYRL----GAVQNEVTLTHPITKYIMACMSADLEVTS-TWVLVGVGVL  1667

Qy   1572 VAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI------VEECASFIP-LEAM  1624
          :: | |  |  |  | |   |:  : |:   | |::|         |:|  ::|  :
Db   1668 AALAAYCLTTGSVVIVGR-------IILSGRPAVIPDREVLYQEFDEMEECASHLPYIEQG  1721

Qy   1625 VAAIDKLK--------STIT----TTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVT  1673
          : | ::|         |  :    : || :||:| ::  |    |  :|:: |    |:
Db   1722 MQLAEQFKQKQKALGLLQTATKQAEAAAP-VVESKWRALEVFWAKHMWNFISGIQYLAGLST  1780

Qy   1674 LPDNPFASCVPAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALG  1733
          |  | | | |     | | :|    :|:| |||:  | |||: || | | ||| |  | |
Db   1781 LPGNPAIASLMAFTASITSPLTTQNTLLFNILGGWVAQLAPPSAASAFVGAGIAGAAVG  1840

Qy   1734 TWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVVSAFNPAAGVVGVL  1790
          :  |:|    ||: | ||: ||::|| ||:  ::| | | |:|| :|| |   |  |||
Db   1841 ---SIGLGKVLVDILAGYGAVGAGLVAFKVMSGEMPSTEDLVNLLPAILSPGALVVGVV  1897

Qy   1791 SACAMFALTTAGPD----HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWS  1846
          || |:| |:| |     ||| || : ||: |::||| |:| |:|:| ||  ||||:| |
Db   1898 --CAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAARVTQILSSLT---  1952

Qy   1847 VISACIRWLHITPTEDDCGLI--AWGLEIWQYVCNFFVICFNVLKAGVQS--MVNIPGCP  1901
          |:| : | | || :| | |    |||:|:|| | ::| | :|||||| :    ||:|| |
Db   1953 -ITQLLKRLHQWINEDCSTPCSGSWLKDVWDWICT----VLSDFKTWLQSKLLPRLPGLP  2007

Qy   1902 FYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARL  1961
          | |||:|||| | ||| | | ||||   :  |  :  :  :|   |:| |  |  || |
Db   2008 FLSCQRGYKGVWRGDGIMQTTCPCGAQITGHVQNGSMRIV-GPKTCSNTWHGTFPINAYT  2066

Qy   1962 CGSARPDPT-DWTSLVVNYGVRDYCYEKMGDHIFVTAVSSPNV-CFTQVP----  2010
          ||:||  |    |  ||:|  ||| | :|  | | |  |   |  | :|
Db   2067 TGPCTPSPAPNYSRALWRVAAEEYVEVTRVGDFHYVTGMTTDNVKCPCQVPAPEFFTEVD  2126

Qy   2011 ----------------PTLRAAVAVD--------GVQVQCYLGEPKTPWTTSACCYGPDGKGKT  2050
          | ||  | |  | :|| |    |:| ::| |
Db   2127 GVRLHRYAPVCKPLLREEVVFQVGLNQYLVGSQLPC---EPE------PDVAVLT  2172

Qy   2051 VKL--PPRVDGHTPGVRMQLNLRDALETND---------CNSTNNTPSDBAAVSALV  2096
          | |    | | : || | | ::| |  |            || |  |||  |  ::|:|
Db   2173 SMLTDPSHITAETAKRRLARGSPPSLASSSAQSLAPSLKATCTTHHDSPDADLIEANLL  2232
```

```
A;Cross-references: UNIPROT:P26663; EMBL:M58335; NID:g329770; PIDN:AAA72945.1; PID:g3297;
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: nonstructural protein NS3 #status predicted <NS3>
F;1230-1237/Region: hepacivirin #status predicted
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,224

Query Match        20.9%;  Score 3218;  DB 1;  Length 3010;
Best Local Similarity  29.8%;  Pred. No. 9.8e-19;
Matches  968;  Conservative 448;  Mismatches 1192;  Indels 642;  Gaps 116;

Qy    8 TSPVPADRTRKNKQTQASYPVSIK-----------TSVERGQRAKRKVQRDARPR  51
Db    3 TNPKPQRKTKRNTNRR----PQDVKFPGGGQIVGGVYLLPRRGPRLGVRAPRKTSERSQPR  59

Qy   52 NYK-----IAGIHDGLQTLAQAALP----------AHGWGRQDPRHKSRN  86
Db   60 GRRQPIPKARRPBG-RTWAQPGVPWPLYGNEGLGWAGWLLSPRGSRPSWGPTDPRRRSRN 118

Qy   87 LGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLF 143
Db  119 LGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNWATGNLPGCSFSIF 177

Qy  144 VVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVICADE---- 198
Db  178 LLALLSCLTTPASAYEVHNV-SGIYHVTNDCSNASIVYEAADLIMHTPGCVPCVREGNSS 236

Qy  199 -CWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLVRHW 257
Db  237 RCWVALTPTLA-ARNVTIPTTIRRHVDLLVGAAAFCSAMYGVGDLCGSVFLVSQLFTFSP 295

Qy  258 LIHIDLNETGTCYLEVPTGIDPGFLG--FIGW-MAGKVEAVIFLKLASQVPYAIATM 312
Db  296 RRHVTLQDCN-----CSIYPGHVSGHRMAWDMMMNWSPTTALVVSQLL-RIPQAVVDM 347

Qy  313 FSSVHYLAVGALIYYASRGKWYQLLLALMLY------IEATSGNPI------ 352
Db  348 VAGAHWGVLAGLAYYSSMAGNWAKVLIVMLLFAGVDGDTHVTGGAQAKTTNRLVSMFASGP 407

Qy  353 --------------------RVPTGCSIAEF------CSPLMIPCP-- 372
Db  408 SQKIQLINTNGSWHINRTALNCNDSLQTGFLAALFYTHSFNSSGCPERMAQCRTIDKFDQ 467

Qy  373 ---------------CHSY----LSENVSEVICYSPKWTRPITLEYNNSISW 405
Db  468 GWGPITYAESSRSDQRPYCWHYPPPQCTIVPASEVCGPVYCFTPS--PVVVGTTDR-- 521

Qy  406 YPYTIP---GARGCMVKFKNNT-----WGCCRIRNVPSY--------CTMG-- 440
Db  522 --FGVPTYRWGENETDVLLLNNTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGGVGNNTL 579

Qy  441 ---TDAVWNDTRNTYEACGVTPWLT------TAWHNGSALKLAILQ--YPGSKE-- 483
Db  580 TCPTDCFRKHPEATYTKCGSGPWLTPRCMVDYPYRLWHYPCTVNFTIFKVRMYVGGVEHR 639

Qy  484 MFKPHNWMSG-HLYFEGSDTPIVYFYDPVNSTLPPERWARLPGTPPVVRGSWLQVPQGF 542
Db  640 LNAACNWTRGERCDLEDRDRP-----ELSPLLSTTEWQVLPCS------- 678

Qy  543 YSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAYLC 602
Db  679 FTTLPALSTGLIHLHQNIVDVQLYLYGIGSAVVSFAIKWEYVLLLFLLLADARVCACLWMM 738
```

```
Qy 2097 FKQEL------RRTNQLL-----EAISAGVDTTKLPAPSIEEVVVRKRQPRARTGSLTL 2144
Db 2233 WRQEMGGNITRVESENKVVILDSFDPIRAVEDEREISVPA-EILRKPRKFPPALPIWAR 2290

Qy 2145 P----PP-----PRSVPGV-SCPESLQRSDPLEGPS-NLPPSPP------VLQLAMP 2184
Db 2291 PDYNPPLLESWKDPDYVPPVHGCPL-------PSTKAPPIPPPRRKRTVVLTESTV 2340

Qy 2185 MPLLGAGECNPFTAIGCAMTETG----GGPD----------DLPSY---PP----KKEVSEW 2225
Db 2341 SSALAELATKTFGSSGSGSSAVDSGTATGPPDQASDDGDKGSDVESYSSMPPLEGEPGDPDL 2400

Qy 2226 SDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTWTD-VI 2284
Db 2401 SDGSWST-----VSG------EAGEDVVCCSMSYTWTGALI 2430

Qy 2285 SFKTASKVLSATRAITSGFLKQRSLVVTEPRDAELRKQKVTINRQPLFPPSYHKQVRLA 2344
Db 2431 TPCAAEESKLPINPLSNSLLRRHSMVYSTTSRSASLRQKKVTFDRLQVLDDHYRDVLKEM 2490

Qy 2345 KEKASKVVGVMNDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDL---QKCVEA 2400
Db 2491 KAKASTVKARLLSIEEACKLTPPHSAKSKF-GYGAKDVRSLSSR-AVNHIRSVWEDLLED 2548

Qy 2401 GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV 2460
Db 2549 TETPI---DTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVSTLPQAV 2605

Qy 2461 MGDAYGF-VDPRTRVKRLLSMMSPD--AVGATCDTVCFDSTITPEDIMVETDIYSAAKLS 2517
Db 2606 MGPSYGFGPQYSPGQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTENDIRTEESIYQCCDLA 2665

Qy 2518 DQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQA 2577
Db 2666 PEARQAIRSLTERLYVGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKATAACRAA 2725

Qy 2578 GMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSC 2637
Db 2726 KLQDCTMLVNGDDLVVICESAGTQDEAAALRAFTEAMTRYSAPPGDPPQPEYDLELITSC 2785

Qy 2638 SSNVTSGITKSGKPYYFLTRDDRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLA 2697
Db 2786 SSNVSVAHDASGKRVYLTRDPTTPLARAAWETVRHTPVNSWLGNIIMYAPTLWARMILM 2845

Qy 2698 VHFMEQMLFEDKLPETVTFPDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVS 2757
Db 2846 THFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLHSYSPGEINRVA 2905

Qy 2758 QSLTDMTMPPLRAWRKKARAVLASAKRRGGANAKLARPLL-WHATSR----PLP-----D 2807
Db 2906 SCLRKLGGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVKTKLKLTPIPAASQLD 2965

Qy 2808 LDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAI 2863
Db 2966 LSGWFVAGYN--------GDIYHSLSRARPRW------FMLCLLLLSGVGVGI 3004
```

RESULT 7

GNWVTC

genome polyprotein - hepatitis C virus

N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: A38465
R;Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.;
J. Virol. 65, 1105-1113, 1991
A;Title: Structure and organization of the hepatitis C virus genome isolated from human
A;Reference number: A38465; MUID:91140698; PMID:1847440
A;Accession: A38465
A;Molecule type: genomic RNA
A;Residues: 1-3010 <TAK>

```
Qy  603  YLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAALLG  662
Db  739  LL------------------------------IAQAEAALENLVV---------------  761

Qy  663  FVPMAAGLPLTFFVAAAAQPDYDWWVR-LLVAGLVLWAGRNRGHRIALLVGPWPLVALL   721
Db  762  ----AHGI-LSFLVFFCAA------WYIKGRLVPGATY--------               798

Qy  722  TLLHLVTPASAFDTRIIGGLTIPPVVALVMSRFGFFAHLLPRCALVNSYLWQRWENWFW   781
Db  799  LLLALLPPRAYAMDREMAASCCGAVFGLVLLTLSPYYKVFLAR------LIW------WLQ 847

Qy  782  NVTLRPERFFLVLVCFP------GATYDALVTFCVCHVALLC-LTSSAASFFG----TDSR  831
Db  848  YFTTRAEADLHVWI--PPLNARGGRDAIILLMCAVHPELIFDITKLLIAIGPLMVLQAG   905

Qy  832  VRAHRMLVRL-GKCHAWY------SHYVLKFFLLVFGENGVFPYKHLHGDVLPNDFASK   883
Db  906  ITRVPYFVRAQGLIHACMLVRKVAGGHYVQMAFMKLGALTGTYIYNHL------------  953

Qy  884  LPLQ----------EPFPPEGKARVYRNEGRRLACGDTVDGLPVVARLGDLVFAG       929
Db  954  TPLRDWPRAGLRDLAVAVEPVVFSDMETKIITWGADTAACGDIIIGLPVSARRGKEILLG 1013

Qy  930  LAMPPD------GWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSYM  984
Db 1014  ---PADSLEGRGLRLLAPITAYSQQTRGLLGCIITSLTGRDKNQVEGEVQVVSTATQSFL 1070

Qy  985  GFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN-DQDI--YQPPCGAGSLTRCSCGE 1041
Db 1071  ATCVNGVCWTVYHGAGSKTLAAPKG---PITQMYTNVDQDLVGWPKPPGARSLTPCTCGS 1127

Qy 1042  TKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAA---R 1098
Db 1128  SDLYLVTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPFGHAVGIFRAAVCTR 1187

Qy 1099  NSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSY 1158
Db 1188  GVAKAVDFVPVESMETTMRSPVFTDNS---SPPAVPQSFQVAHLHAPTGSGKSTKVPAAY 1244

Qy 1159  MQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYYNGKCTNTGASLTYSTYGMYLT-GA 1217
Db 1245  AAQGYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGAPVTYSTYGKFLADGG 1304

Qy 1218  CSRN-YDVIICDECHATDATVLGIGKVLTEAPSKNVRLVVLATATPPGVLPTPHANITE  1276
Db 1305  CSGGAYDIIICDECHSTDSTTILGIGTVLDQAETAGARLVVLATATPPGSVTVPHNIEE  1364

Qy 1277  IQLTDEGTIPHGKKIKEENLKKGRHLIFEATKKHCDELANELARKGITAVSYYRGCDIS  1336
Db 1365  VALSNTGEIPFYGKAIPIEAIRGGRHLIFCHSKKKCDELAAKLSGLGINAVAYYRGLDVS 1424

Qy 1337  KIPE-GDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVK 1395
Db 1425  VIPTIGDVVVVATDALMTGYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSR 1484

Qy 1396  GQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTERQTILDTYRT 1455
Db 1485  SQRRGRTGRGRRGIYRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAETSVRLRAYLN 1544

Qy 1456  QPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLCHQYGYAAPND   1514
Db 1545  TPGLPVCQDHLEFWESVFTGLTHIDAHFLSQTKQAGDNFPYLVAYQATVCARAQAPPSW  1604

Qy 1515  APRWQG-ARLGKK-----PCGVLWRLDGDADACPGPEPSEV-------TRYQM-CFT--EVNT 1560
Db 1605  DQMWKCLIRLKPTLHGPTPLLYRL-------GAVQNEVTLTHPITKYIMACMSADLEVVT 1657

Qy 1561  SGTAALAVGVGMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEE------BIVEE   1614
Db 1658  S-TWVLVGGVLAALAAYCLLTGSVVIVGR------IILSGRPAIVPDRELLLYQEFDEMEE 1710

Qy 1615  CASFIP-LEAMVAAIDKLK------STIT----TTSPFLTETALEKLNTFIGPHAATIL  1662


Db 1711  CASHLPYIEQGMQLAEQFKQKALGLLQTATKQAEAAAP-VVESKWRALETFWAKHMWNFI 1769
Qy 1663  AIIEYCCGLVTLLPDNPFAASCVPAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAF 1722

Db 1770  SGIQYLAGLSTLPGNPAIASLMAFTASITSPLTTQSTLLFNILGGWVAAQLAPPSAASAF 1829
Qy 1723  MMAGAGTALGTWTSVGF---VFDMLGGYAAASTACLTFKCLMGEWPTMDQLAGLVVSA   1779

Db 1830  VGAGIAGAAVG---SIGLGKVLVDILAGYGAGVAGALVAFKVMSGEMPSTEDLVNLLPAI 1886
Qy 1780  FNPAAGVGVLSACAMFALTTAGPD---HWPNRLLITMLARSNTVCNEYFIATRDIRRKI  1835

Db 1887  LSPGALVGVV--CAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAARV  1944
Qy 1836  LGILEASTPWSVISACIRWLHTPTEDDCGLI---AWGLEIWQYVCNFFVICFNVLKAGVQ 1892

Db 1945  TQILSSLT----ITQLLKRLHQWINEDCSTPCSGSWLRDVWDWICT----VLTDFKTWLQ 1996
Qy 1893  S--MVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNY 1950

Db 1997  SKLLPQLPGVPFFSCQRGYKGVWRGDGIMQTTCPCGAQITGHVKNGSMRIV-GPKTCSNT 2055
Qy 1951  WRGAVPVNARLCGSARPDT-DWTSLVVNYGVRDYCYEKMGDHIFVTAVSSPNV-CFTQ   2008

Db 2056  WHGTFPINAYTTGPCTPSPAPNYSRALWRVAAEBYVEVTRVGDFHYVTGMTDNVKCPCQ  2115
Qy 2009  VP--------PTLRAAAVD---------GVQVQCYLGEPKTPWTTSA              2039

Db 2116  VPAPEFFSEVDGVRLHRYAPACRPLLREEVTFQVGLNQYLVGSQLPC---EPE------  2165
Qy 2040  CCYGPDGKGKTVKL--PFRVDGHTPGVRMQLNLRDALETND----------CNSTNNT   2085

Db 2166  ----PDVAVLTSMLTDPSHITAETAKRRLARGSPPSLASSASQLSAPSLKATCTTHHVS  2221
Qy 2086  PSDEAAVSALVFKQEL-------RRTNQLL---EAISAGVDTTKLPAPSIEEVVRKR    2133

Db 2222  PDADLIEANLLWRQEMGGNITRVESENKVVVLDSFDPLRAEEDEREVSVPA--EILRKSK 2279
Qy 2134  QPRARTGSLTLP-----PP-----PRSVPGV--SCPESLQRSDPLEGPSNLPP-------  2174

Db 2280  KPPAAMPIWARPDYNPPLLESWKDPDYVPPVVHGCP-----LPPIKAPPIPPPRRKRTVV 2334
Qy 2175  -------SPPVLQLAMPMPLLGGAGECNPP-------TAIGCAMTETGGGPDDLPSY---PP- 2218

Db 2335  LTESSVSSALAELA--TKTFGSSESSAVDSGTATALPDQASDDGDKGSDVESYSSMPPLE 2392
Qy 2219  -KKEVSEWSDESWSTATTASSVVTGPPYPKIRGKDSTQSAPAKRPTKKLGKSEFSCCSMS 2277

Db 2393  GEPGDPDLSDGSWSTVSEEAS--------EDVVCCSMS                       2422
Qy 2278  YTWTD-VISFKTASKVLSATRAITSGFLKQRSLVYTEPRDAELRKQKVTINRQPLFPPS  2336

Db 2423  YTWTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVLDDH 2482
Qy 2337  YHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQK 2396

Db 2483  YRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKF-GYGAKDVRN-LSSKAV-----  2535
Qy 2397  CVEAGEIPSHYR----------QTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEK  2446

Db 2536  ----NHIHSVWKDLLEDTVPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEK  2591
Qy 2447  MYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMW--SPDAVGATCDTVCFDSTITPED 2503

Db 2592  MALYDVVSTLPQVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTEND  2651
Qy 2504  IMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNS 2563

Db 2652  IRVEESIYQCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCCGYRRCRASGVLTTSCGNT 2711
Qy 2564  LTCWLKVNAAAEQAGMKNPRFLICGGDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDC 2623
```

```
Qy      8 TSVPAPRTRKNKQTQASYPVSIK-----------------------------TSVERGQRAKRKVQRDARPR   51
Db      3 TNPKP---QKKNKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR               59
Qy     52 NYK----IAGIHDGLQTLAQAALP-------------------------AHGWGRQDPRHKSRN           86
Db     60 GRRQPIPKARRPEG-RTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRN              118
Qy     87 LGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVhLF              143
Db    119 LGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLPGCSFSIF              177
Qy    144 VVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC----AD               197
Db    178 LLALLSCLTVPASAYQVRN-STGLYHVTNDCPNSSIVEAADAILHTPGCVPCVREGNAS               236
Qy    198 ECWVPANPYI-SHPSNWTGTDSFPLADHIDFVMGALVTCDALDIGELGCACVLVGDWLV-              254
Db    237 RCWVAMTPVATRDGKLPATQ--LRRHIDLLVGSATLCSALVYGDLCGSVFLVGQLFTFS               294
Qy    255 --RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGW-------------------------              309
Db    295 PRRHW------TTQGCNCSIYPGHITGHR-MAWDMMNWSPTTALVMAQLL-RIPQAI                 344
Qy    310 ATMFSSVHYLAVGALIYASRGKWYQLLLALMLY--IEA                                    346
Db    345 LDMIAGAHWGVLAGIAYFSMVGNWAKVLVVLLLFAGVDAETHVTGGSAGHTVSGFVSLLA              404
Qy    347 --------------------------------------TSGNPIRVPTGCSIAE                   362
Db    405 PGAKQNVQLINTNGSWHLNSTALNCNDSLNTGWLAGLFYHHKFNSSGCPERLAS------              458
Qy    363 FCSPL----------------------MIPCPCHSYLSENV-SEVICYSPKWTRPI                  395
Db    459 -CRPLTDFDQGWGPISYANGSGPDQRYCWHYPPKPCGIVPAKSVCGPVYCFTPS---PV               514
Qy    396 TLEYNN-----SISWYPYTIPGARGCMVKFKNNT------WGCCRIRNVPSY------                436
Db    515 VVGTTDRSGAPTYSW------GENDTDVFVLNNTRPPLGNWFGCTWMNSTGFTKVCGAPP              568
Qy    437 CTMG------TDAVWNDTRNTYRACGVTPWLT------TAWHNGSALKLAILQ                     477
Db    569 CVIGGAGNNTLHCPTDCFRKHPDATYSRCGSGPWITPRCLVDYPYRLWHYPCTINYTFK               628
Qy    478 ---------YPGSKE--MEKPHNWMSG--HLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVV       531
Db    629 IRMYVGGVEHRLEAACWTRGERCDLEDRDR----SELSPLLLTTTQWQVLPCS-----                678
Qy    532 RGSWLQVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLLLGLC               591
Db    679 ----------FTTLPALSTGLIHLHQNIVDVQYLYGVGSSIASWAIKWEYVVLFLLLA                727
Qy    592 GSKYLILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICYCYL             651
Db    728 DARVCSCLWMMLL-----------------------ISQAEAALENLVI------                   753
Qy    652 RCRLRYAALLGFVPMAAGLP--LTFFVAAAAQPDYDWWVRLLVAGLVLWAGNRGHRIA                709
Db    754 ---LNAASLAG---THGLVSFLVFFCFA-----WYLK---GKWVPGAVY                         788
Qy    710 LLVGPWPLVALLTLLHLVTPASAFDTEI---IGGLTIPPVVALVVMSRFGFFAHLLPRCA              766
Db    789 TFYGWMPL--LILLLALPQRAYALDTEVAASCCGVVL--VGLMALTLSPYYKRYISWCL               843
Qy    767 LVNSYLMQRWEN---WFWNVTLRPERPFLVLVCFPGGATYDALVTFCVCHVALLC-LTSS              821
Db    844 WWIQYFLTRVEAQLHVWIPPLNVRGGRDAVILL--------MCAVHPTLVFDITKL                  891
Qy    822 AASFFGTDSRVRAHRMLY----------------------RLGKCHAWYSHYVLKFPLLVFGEN          863
Db    892 LLAVFGPLWILQASLLKVPYFVRVQGLLRFCALARKMIG----GHVQMVIKLGALT                  945
Qy    864 GVFFYKHLH--GDVLPNDFASKLPLQEPFPFPEGKARVYRNEGRRLACCGDTVDGLPVVAR             921
```

```
Db   2712 LTCYLKASAACRAAKLQDCTMLVNGDDLVICESAGTQEDAASLRVFTEAMTRYSAPPGD              2771
Qy   2624 VPQPKYSLEELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYL             2683
Db   2772 PPQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNI             2831
Qy   2684 IHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAF             2743
Db   2832 IMYAPTIWARMLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAF             2891
Qy   2744 SVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WHATS            2802
Db   2892 SLHSYSSPGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVKT           2951
Qy   2803 R-----PLP-----DLDKTSVARYTTFNVCDVSPEGDVFITPQRRLQKFLVKYLAVIVFA            2853
Db   2952 KLKLTPIPAASRLDLSGWFVAGYS--------GGDIYHSLSRARPRW------FMLC               2994
Qy   2854 LGLIAVGLAI                                                             2863
Db   2995 LLLLSVGVGI                                                             3004
```

RESULT 8
GNWVC3
genome polyprotein - hepatitis C virus (strain HCV-1)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C;Accession: A39166; PQ0403; PQ0404
R;Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Coi
Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
A;Title: Genetic organization and diversity of the hepatitis C virus.
A;Reference number: A39166; MUID:91172826; PMID:1848704
A;Accession: A39166
A;Molecule type: mRNA
A;Residues: 1-3011 <CHO>
A;Cross-references: UNIPROT:P26664; GB:M62321; NID:g329873; PIDN:AAA45676.1; PID:g329874
R;Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.L
J. Gen. Virol. 73, 1131-1141, 1992
A;Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e
A;Reference number: PQ0393; MUID:92268871; PMID:1316939
A;Accession: PQ0403
A;Molecule type: genomic RNA
A;Residues: 1577-1633 <CHA>
A;Cross-reference: DDBJ:D10128
A;Experimental source: isolates E-b16
A;Accession: PQ0404
A;Molecule type: genomic RNA
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1577-1633 <CH2>
A;Experimental source: isolates E-b17
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
F;1-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1319/Region: nucleotide-binding motif B
F;1313-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2077,22

Query Match   20.8%;   Score 3204.5; DB 1; Length 3011;
Best Local Similarity   29.7%;   Pred. No. 6.8e-191;
Matches 962; Conservative 442; Mismatches 1215; Indels 621; Gaps 110;

Db   1998 KLMPQLPGIPFVSCQRGYKGVWRVDGIMHTRCHCGAEITGHVKNGTMRIV-GPRTCRNMW 2056
Qy   1952 RGAVPVNARLCGSARPDT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFTQV 2009
Db   2057 SGTFPINAYTTGPCTPLPAPNTPALMRVSAEEYVEIRQVGDFHYVTGMTTDNLKCPCQV 2116
Qy   2010 P--------------------PTLRAAVAVD----------GVOVQCYLGEPKTPWTTSAC 2040
Db   2117 PSPEFFTELDGVRLHRFAPPCKPLLREEVSFRVGLHBYPVGSQLPC---EPE------- 2165
Qy   2041 CYGPDGKGKTVKJ--PFRVDGHTPGVRM-------------QLNLRDALETNDCNSTNN 2084
Db   2166 ---PDVAVLTSMLTDPSHITAEAAGRRLARGSPPSVASSSASQLSAPSLKAT--CTANHD 2220
Qy   2085 TPSDEAVSALVFKQEL-------RRTNQLL-------EAISAGVDTTKLPAPSIEEVVRRK 2132
Db   2221 SPDAELIEANLLWRQEMGGNITRVESENKVVILDSFDPLVAEEDEREISVPA--EILRKS 2278
Qy   2133 RQFRARTGSLTLPPPPRSVPGVSCP--ESLQRSD---------PLEGPSNLPPSPP----- 2177
Db   2279 RRF------AQALPVWAR--PDYNPPLVETWKKPDYEPPVHGCPLPPKSPPVPPPRRKR 2331
Qy   2178 --VL---QLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPS-YPPKKEVSEW----- 2225
Db   2332 TVVLTESTLSTALAELATRSFGSSSTSGITGDNTTTSSEPAPSGCPPDSDABSYSSMPPL 2391
Qy   2226 -----------SDESWSTATTASSYVTGPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSM 2276
Db   2392 EGEPGDPDLSDGSWSTVSSEAN----------------ABDVVCCSM 2422
Qy   2277 SYTWTDVISFKTASKVLS-ATRAITSGFLKQRSLVVVTEPRDAELRKQKVTINRQPLFPP 2335
Db   2423 SYSWTGALVTPCAABEQKLPINALSNSLLRHNLVYSTTSRSACQRQKKVTFDRLQVLDS 2482
Qy   2336 SYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDL- 2394
Db   2483 HYQDVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKF-GYGAKDVRC-HARKAVTHIN 2540
Qy   2395 ---QKCVEAGEIPSHYRQTVIVPKEEVVFKVTPQKPTKPPRLISYPHLEMRCVEKMYYGQ 2451
Db   2541 SVWKDLLEDNVTPI---DTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYD 2597
Qy   2452 VAPDVVKAVMGDAYGF-VDPRTRVKRLLSMW--SPDAVGATCDTVCFDSTITPEDIMVET 2508
Db   2598 VVTKLPLAVMGGSSYGFQYSPGQRVEFLVQAWKSKKTPMGFSYDTRCFDSTVTESDIRTEE 2657
Qy   2509 DIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVVTTSSNSLTCWL 2558
Db   2658 AIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGVLTTSCGNTLTCYI 2717
Qy   2569 KVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPK 2628
Db   2718 KARAACRAAGLGQDCTMLVCGDDLVVICESAGVQEDAASLRAFTEAMTRYSAPPGDPPQPE 2777
Qy   2629 YSLEELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYP 2688
Db   2778 YDLELITSCSSNVSVAHDGAGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMFAP 2837
Qy   2689 CLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLLPSIIAGVHGIBAFSVVRY 2748
Db   2838 TLWARMILMTHFFSVLLARDQLEQALDCEIYGACYSIEPLDLPIIQRLHGLSAFSLHSY 2897
Qy   2749 TNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRGGAHAKLARFLL-WHATSR---- 2803
Db   2898 SPGEINRVAACLRKLGVPPLRAWRHRARSVRARLLARGGRAAICGKYLFNWAVRTKLKLT 2957
Qy   2804 PLPDLDKTSVARYTTFNYCDVVSPEGDVFITPQRRLQKFLVKYLAVIVPALGLIAVGLAI 2863
Db   2958 PIAAAGQLDLSGWFTAGYSG-----GDIYHSVSHARPRW-------IMFCLLLLAAGVGI 3005

Db   946 GTYVYNHLTPLRDWAHNGLRDLAVAVEPVVPSQMETKLIITWGADTAACGDIINGLPVSAR 1005
Qy   922 LGDLVFAGLA--MPPDGWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSL   979
Db   1006 RGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTA 1065
Qy   980 ATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDI--YQPPCGAGSLTRC 1037
Db   1066 AQTFLATCINGVCWTVYHGAGTRTIASPKGPV--IQMYTNVDQDLVGWPAPQGSRSLTPC 1123
Qy   1038 SCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAA 1097
Db   1124 TCGSSDIYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAA 1183
Qy   1098 ---RNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILIAPTGSGKSTKL 1154
Db   1184 VCTRGVAKAVDFIPVENLETTMRSPVFTDNS---SPPVVPQSFQVAHLHAPTGSGKSTKV 1240
Qy   1155 PLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYL 1214
Db   1241 PAAYAAQGYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFL 1300
Qy   1215 T-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHA 1272
Db   1301 ADGGCSGGAYDIICDECHSTDATSLIGIGTVLDQAETAGARLVVLATATPPGSVTVVPHP 1360
Qy   1273 NITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELARKGITAVSYYRG 1332
Db   1361 NIEEVALSTTGEIPFYGKAIPLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRG 1420
Qy   1333 CDISKIP--EGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVS 1391
Db   1421 LDVSVIPTSGDVVVVATDALMTGYTGDFDSVIDCNTCVTQTFVDFSLDPTFTIETTLPQD 1480
Qy   1392 AIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAPDAAKAWYGLSSTEAQTILD 1451
Db   1481 AVSRTQRRGRTGRGKPGIYRFVAPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLR 1540
Qy   1452 TYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLQLCHQYGYA 1510
Db   1541 AYMNTGPLPVCQDHLEFWEGVFTGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAP 1600
Qy   1511 APNDAPRWQG-ARLGKK----PCGVLWRLDGADACPGPEPSE-----VTRYQM-CFT--- 1556
Db   1601 PPSWDQMWKCLIRLKPTLHGPTPLLYRL----------GAVQNEITLTHPVTKYIMTCMSADL 1653
Qy   1557 EVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI------ 1611
Db   1654 EVVTS-TWVLVGGVLAGLSTLPGNPAIASLMAFTAAVTSPLTTSQTLLFNILGGWVAAQLAAPGA 1706
Qy   1612 -VEECASFIPL--EAMVAA------IDKLKSTITTTS--------PFTLETALEKLNTFLGPHA 1658
Db   1707 EMEECSQHLPYIEQGMMLAEQFKQKALGLLQTASRQAEVIAP-AVQTNWQKLETFWAKHM 1765
Qy   1659 ATILAIIEYCCGLVTLPDNPPASCVAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARG 1718
Db   1766 WNFISGIQYLAGLSTLPGNPAIASLMAFTAAVTSPLTTSQTLLFNILGGWVAAQLAAPGA 1825
Qy   1719 ALAFMMAGAGTALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGL 1775
Db   1826 ATAFVGAGLAGAAIG---SVGLGKVLIDILAGYGAGVAGALVAFKIMSGEVPSTEDLVNL 1882
Qy   1776 VYSAFNPAAGVGVGVLSACAMFALTTAGPD----HWPNRLLTMLARSNTVCNEYFIATRDI 1831
Db   1883 LPAILSPGALLVGVV--CAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHVYPESDA 1940
Qy   1832 RRKILGLILEASTPWSVISACIRWLHTPTEDDCGLIAWGLEIWQYYCNFFVICFNVLKAGV 1891
Db   1941 AARVTAILSSLLTVTQLLRRLHQWISSECTTPCS-GSWLRDIWDWICEVLSDFKTWLKA-- 1997
Qy   1892 QSMVNIPGCPFYSCQKGYKGPWIGSSMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYW 1951

RESULT 9
S18030

```
genome polyprotein - hepatitis C virus (isolate JK1)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
  protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
A;Variety: isolate JK1
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: S18030, S33570; A48332, S18029
R;Honda, M.; Kaneko, S.; Kobayashi, K.; Murakami, S.
A;Description: A whole genome of hepatitis C virus cDNA was isolated from a single patie
  submitted to the EMBL Data Library, September 1991
A;Reference number: S18028
A;Accession: S18030
A;Molecule type: genomic RNA
A;Residues: 1-3010 <HON>
A;Cross-references: UNIPROT:Q68949; EMBL:X61596; NID:g59478; PIDN:CAA43793.1; PID:g59479
A;Experimental source: isolate JK1 from an individual
R;Honda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
A;Title: Sequence analysis of putative structural regions of hepatitis C virus isolated
  Arch. Virol. 128, 163-169, 1993
A;Reference number: A48332; MUID:93119270; PMID:8380322
A;Accession: S33570
A;Molecule type: genomic RNA
A;Residues: 1-547,'T',549-621,'V',623-624,'S',626-652,'DL',655-761,'T',763-782 <HOW>
A;Cross-references: EMBL:X61591
A;Note: this sequence is inconsistent with the nucleotide translation
A;Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue 320
  as Trp, and TTC for residue 771 as Ser
A;Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBIP:121748)
C;Superfamily: hepatitis C virus genome polyprotein
C;Keyword: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Region: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1615-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate (As

Query Match             20.6%;  Score 3161;  DB 1;  Length 3010;
Best Local Similarity   29.3%;  Pred. No. 3.5e-188;
Matches  947;  Conservative 453;  Mismatches 1231;  Indels 596;  Gaps 106;
```

```
Qy     8  TSPVPAPRTRKNKQTQASYPVSIK------------------TSVERGQRAKRKVQRDARPR   51
             ||:|  ::   |||                         :|| :||
Db     3  TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR   59

Qy    52  NYKIAGIHDGLQTLAQAALP--------------------AHG----WGRQD          79
             :  :||     :|                            :
Db    60  GRR---------QPIPKARQPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPYGSRPRWGPTD 111

Qy    80  PRHKSRNLGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPVCGIVRLLEDGVNWATG---  136
          || |||||| ::| | :|  |  || |||| |  :   |   | ||||||| |||
Db   112  PRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLP 170

Qy   137  WFGVHLFVVCLLS-LACPCSGARVTDPPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC 195
            | : | :| ||  | | | |  |     :   |:  | | :|| |    | :::|| |
Db   171  GCSFSIFLLALLSCLIVPVSTVEVRNV-SGVHVHTNDCSNSSIVYEAADMIMHTPGCVPC  229

Qy   196  ADE-----CWPANPYIHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVG    250
           :         |  :| :  | |||  ||   | :| ||  |  |  | |:|| | ::|
Db   230  VREGNSSRCWVALTPTLAARNSSIPTTT-IRRHVDLLVGAAALCSAMYVGDLCGSVFIVS  288

Qy   251  DMLV--RHWLIHIDLNETGTCYLEVPTGIDPGFLG--FIGW--MAGKVEAVIFLTKLA   302
             ::    |  ||:|:     |   | :| :     ::|:
Db   289  QLFTFSPRRYETVQDCN---CSLY---PGHVSGHRMAWDMMNWSPTTALVVSQLL       338

Qy   303  SQVPYAIATMFSSVHYLAVGALLIYYASRGKWYQLLLALMLYIEATSGNPIRV---PTGC  358
          :::|||:  :| :| | : |  : :|  | ||  ||| | :|:
Db   339  -RIPQAVVDMVVGAHWGVLAGLAYYSMVGNWAKVLIVMLLFAGVDGTTYVSGHASQTTR   397

Qy   359  SIAEFCSP--------------------LMIPCCHSYLSENVSEVICYSPKWTR        393
           :|| ||                      |: |  : |: :  :|: | |
Db   398  RVASFFSPGSAQKIQLVNTNGSWHINRTALNCNESINTGFFAALFVKKFNSSGCSERMA   457

Qy   394  ------PITLEYNNSISWYPY-------------------TIPGARGC               416
                    | :  : |:|                        ||  | |
Db   458  SCRPIDRFAQGWGPITHAESRSSDQRPYCWHYAPQPCGIVPALQVCGPVYCFTPSPVVVG  517

Qy   417  -------MVKFKNT------WGCCRIRNVPSY------CTMG                    440
                  :|| ::      | |:| :: :         :|
Db   518  TTDRFGVPTYNWGDNETDVLLLNNTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGGAGNN  577

Qy   441  -----TDAVWNDTRNTYEACGVTPWLT-----TAWHNGSALKLAILQ--YPGSKE       483
                  || |  |:  :|||| ||||      || |  : | ||::
Db   578  TLTCPTDCFRKHPEATVTKCGSGPWLTPRCMVDYPYRLWHYPCTFNFTIFKIRMYVGGVE  637

Qy   484  --MFKPHNWMSG-HLYPEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ  540
            : :  :| :      |  :|:| : |:||                 : :: |
Db   638  HRLNAACNWTRGERCNIEDRDR------SELSPLLLSTTEWQILPCS------         678

Qy   541  GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLLAY   600
           : |  : :|  |: : : : :|  | | ||| || :| |: :| | ||| ||| || ||
Db   679  --FTTLPALSTGLIHLHQNIVDVQYLYGVGSAVVSIVIKWEYVLLLFLLLADARVCACLW  736

Qy   601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL  660
            : :    :       :       : :|     :|   |:| |  |
Db   737  MMLL-----------------------IAQAEAALENLVV----------LNAASV     759

Qy   661  LGFVPMAAGLPLTFFVAAAAQPDYDWWVR-LLVAGLVLWAGRNRGHRIALLVGPWPLVA   719
          |||  |||||| | | |   :|  ||::   : :|:: :  | | | || ||| ||:
Db   760  AG---AHGI-LSFLVFFCAA-----WYIKGRLVPGAAY----------AFYGVWPL--    796

Qy   720  LLTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQRWEN-  778
          || | ||| |||| ||| |||| ||||| | |:||||||||||||||| ::| || |
Db   797  LLLLLALLPPRAYAMDREMAAASCGGAVFVGLALLTLSPHYKAFLARLIWWLQYFITRVEAC 856

Qy   779  ---WFWNVTLRPERFFLVLVCFPGATYDALVTFCVCHVALLC-LTSSAASFFGTDSRVRA  834
             | |:   |  :|:|: :|  |:||| :   |:| | :     :::|| :||:| :| |
Db   857  LQVWVPPLNVRGGRDAILLT----------------CAVHSELIFDITKILLAIFGP--  898

Qy   835  HRMLVRLGKCHAWY------------------SHYVLKFFLLVFGENGVFFYKHLH---G  873
           |: |:| |                       :||::||||||||| | || || |   |
Db   899  -LMVLQAGLTRVPYFVRAQGLLRACMLVRKVAGGHYVQMALMKLIAALTGTYVVDHLTPLR 957

Qy   874  DVLPNDFASKLPLQEPFPFPEGKARVIRNEGRRLACGDTVDGLPVVARLGDLVFAGLAMP  933
          |   :|    :|  |  : : :  : | : :  ||||| : ||| ::|   ||| |   P
Db   958  DWAHSGLRDLVVAVEPVVFSDMETKIITWGADTAACGDIILGLPVSARRGREILLG---P 1014

Qy   934  PD-----GWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSYMGFVC  988
                  ||    || | |     ||:|||| |:|||    :  :: |:|  | :|:|||
Db  1015  ADGFREQGWRLLAPITAYSQQTRGLFGCIVTSLTGRDKNQVEGEAQVVSTATQSFLATCV 1074

Qy   989  DNVLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDI--YQPPCGAGSLTRCSCGETKGYL 1046
          :|| |||::|: | :|| | |   ::: :  :  : :  ::  | |:||:|| | : :||
Db  1075  NGVCWTVTVHGAGSKTLAGPKGPINQMYTNV--DQDLVGWQAPSGAASLTPCTYGSSDLYL 1132

Qy  1047  VTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVGMFTAA--RNSGGS   1103
          ||| :: :  :   |    : || |:| :||||| || |||| ||:   |  ||:|||
Db  1133  VTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIPRAAVCTRGVAKA 1192

Qy  1104  VSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQLLIAPTGSGKSTKLPLSYMQEKY 1163
          |  | :    | |||| :  : |  : |  |   :  || |||||||||| | : :|: |
Db  1193  VDFIPVESMETIMRSPVFTDNS---SPPAVPQTFQVAHLHAPTGSGKSTKVPAAYAAQGY 1249

Qy  1164  EVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYLT-GACSRN- 1221
          :||||||||| | :  :|: | ||::|| :    :  |||: ||||||  || | || :
Db  1250  KVLVLNPSVAATLGFGAYMSKAHGVDPNISTGVRTITTGAPITYSTYGKFLADGGCSGGA 1309

Qy  1222  YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVLATATPPGVIPTPHANITEIQLTD  1281
          |||||||||||| |||||||||| |||||||||||:|||| |:|| | ||  ||| |:|
```

```
Qy   2235 TASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTWTD-VISFKTASKVL 2293
             :: :.: ::                                    :  :.:::::||:  :
Db   2410 EEAS-----------------------------------------EDVACCSMSYTWTGALITPCAAEESK 2439

Qy   2294 SATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQVRLAKEKASKVVG 2353
          ::   :  |||    .|: :. : :::::::  :     :|  : ||.   |  ||||||
Db   2440 LPINPLSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVPDDHYRDVLKEMKAKASTVKA 2499

Qy   2354 VMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDL----QKCVEAGEIPSHYRQ 2409
          .  .::. | .|: :::|.    |.|  ||..      ||.:    .|     :  |   D
Db   2500 KLLSVEEACKLTPPHSARSKF-GYGAKDVRN-LSSKAVNHIHSVWKDLLEDTETPI---D 2554

Qy   2410 TVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAVMGDAYGF-V 2468
          | |   :||| . | :   |  ||| .| | .||.|||  . |:|::|      .:||. .
Db   2555 TTIMAKNEVFCVQPEKGGRKPARLIVFPELGVRVCEKMALYDVVSTLPQAVMGSSYGFQY 2614

Qy   2469 DPRTRVKRLLSMW--SPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQHRAGIHT 2526
            . || . | .|    . :   | |:.||||.|..:|| .|. ||      |..    | 
Db   2615 SPGQRVEFLVNAWKSKKNPMGFAYCTRCFDSTVTESDIRVEESIYQCCDLAPEARQVIRS 2674

Qy   2527 IARQLYAGGPMIAYDRGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMKNPRFLI 2586
          . | ||.|||: .::.  . |||||.||| ||  . ||||.||.  ||    | ::  .| 
Db   2675 LTERLYIGGPLTNSKGQNCGYRRCRASGVLTTNCGNTLTCYLKASAACRAAKLQDCTMLV 2734

Qy   2587 CGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSNVTSGIT 2646
          ||||  ||  ||| : :. .:||| :.|   ||| |   || |  |: |:|||||   :
Db   2735 CGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHD 2794

Qy   2647 KSGKPYPYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHFMEQMLF 2706
           :|| . |.||||| ||  .  :|   | |.: |:||:| :| ||| || : .||     
Db   2795 ASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLWARMILMTHFFSILLA 2854

Qy   2707 EDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSLTDWTMP 2766
            :|  .. . .    | .  |||| ||  :.|: |||.:: .. ||.|:| .   :  |
Db   2855 QEQLEKALGCQLYGATYFIEPLDLPQIIQRLHGLSAFSLHSYSPGEINRVASCLRKLGVP 2914

Qy   2767 PLRAWRKKARAVLASAKRRGGAHAKLARFLL-WHATSR-----PLP-----DLDKTSVARY 2816
          ||| ||: ||:: |. ::.  |  :. .::. |.|.|      || |     :|: :||.|
Db   2915 PLRVWRHRARSVRAKLLSQGGRAATCGKYLFNWAVRTKLKLTPIPAASQLDLSGWFVAGY 2974

Qy   2817 TTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAI 2863
                       .      :| :  :         .::| | .||.||:
Db   2975 S----------GGDIYHSLSRARPRW------FMWCLLLLSVGVGI 3004
```

RESULT 10
GNWVCH
genome polyprotein - hepatitis C virus (strain H)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructural
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
A;Accession: A36814; A41546
R;Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
submitted to GenBank, July 1992
A;Description: Genomic structure of the human prototype strain H of hepatitis C virus: co
A;Reference number: A36814
A;Accession: A36814
A;Molecule type: genomic RNA
A;Residues: 1-3011 <INC>
A;Cross-references: UNIPROT:P27958; GB:M67463; NID:g329737; PIDN:AAA45534.1; PID:g329738
R;Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991
A;Title: Genomic structure of the human prototype strain H of hepatitis C virus: comparis
A;Reference number: A41546; MUID:92052256; PMID:1658800
A;Contents: annotation
A;Note: neither amino acid nor nucleotide sequence is given
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural
F;1-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>

```
Db   1310 YDIIICDECHSTDSTSILGIGTVLDQAETAGARLVLAAATPPGSVTVPHPNIEBVALPN 1369

Qy   1282 EGTIPPHGKKIKEENLKKGRHLIFATKKHCDELANELARKGITAVSYYRGCDISKIP-E 1340
          :| ||. | .|   . ||||||||  .|:|||||::|:|||||||:|||| |:|.:|  |
Db   1370 TGEIPFYGKAIPLETIKGGRHLIFCHSKKKCCDELAAKLSALGVNAVAYYRGLDVSVIPTS 1429

Qy   1341 GDCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRG 1400
          ||.||||||| ||||||||||.||.  :  |||||  ..:| ..|  . ||:|.:.|||
Db   1430 GDVVVVATDALMTGYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTLPQDAVSRSQRRG 1489

Qy   1401 RTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLP 1460
          |||||| |||:. .   ||||| ..|::: | ||| |||.|:  ||:  ||:|. |||||
Db   1490 RTGRGRGGIYRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAVTSVRLRAYLNTPGLP 1549

Qy   1461 AIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLCHQYGYAAPNDAPRWQ 1519
           :   |:|| .:|  .|  .::|:.   | | :. | |: ||| :||| |     .| |
Db   1550 VCQVHLEFWESVFTGLTHIDAHFLSQTKQAGENFPYLVAYQATVCARAQAPPPSWDQMWK 1609

Qy   1520 G-ARLGKK----PCGVLWRLDGADACPGPEPSEV-----TRYQM-CFT---EVNTSGTAA 1565
           . |||       .||:    .. ..|    |      .  | : |       |.  |  
Db   1610 CLIRLKPTLHGPTPLLYRL------GAVQNEVTLTHPITKFIMACMSADLEVVTS-TWV 1661

Qy   1566 LAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI------VEBCASFI 1619
          | .|:..|  |. .. .. :.  .   |.|:|. | : .:.:|        ||.|||.:
Db   1662 LVGGVLAALAAYCLTTGSVVIVGR-------IILSGRPAIIPDREVLYQEFDEMEBCASHL 1715

Qy   1620 P-LEAMVAAIDKLK-----------STITTTSPFTLETALEKLNTFLGPHAATILAIIEYC 1668
          | :|  . |::: |             |:|.: |||  ||.||| :   |||::||:||.|
Db   1716 PYIEQGMQLAEQFKQKALGLLQTASKQAEAAAPVVESKWQALEEAFWAKHMWNFISSGIQYL 1775

Qy   1669 CGLVTLPDNPPASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAA 1728
           |: ||| |:| :.:.||.|  |::|     :.  |.  :.|: | :|| ||||..|||:|
Db   1776 AGLSTLPGNPAIVSLMAFTASITSPLTTQHTLLFNILGGWVAAQLAPPSAASAFVGAGIA 1835

Qy   1729 GTALGTWTSVGF---VFDMLGGYAAASSTACLTFEKCLMGEWPTMDQLAGLVYSAFNPAAG 1785
          |:|.|     |    :. :|.|||| :. : :  : .:.|| |:||:| .|:.     .|
Db   1836 GAAVG---SIGLGKVLVDILAGYGAGVAGALVAFKGMSGEMPSTEDLVNLLPAILSPGAL 1892

Qy   1786 VVGVLSACAMFALTTAGPD---HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEA 1841
          |||| .:|: ||:: ||      . .|| :  |  :..: ..  |||||||:||| ||:|
Db   1893 VVGVV--CAAILRRHVGPGEGAVQMMNRLIAFASSRGNHVSPTHYVPESDAAARVTKILSS 1950

Qy   1842 STPWSVISACIRWLHTPTEDDCGLI---AWGLEIWQYVCNFFVICFNVLKAGVQS--MVN 1896
           || .:: :|.|.|| .  :|.    | ||| | |   ||::.:|.|:|.|.    :|:
Db   1951 LT-----ITQRLRRLHQWINEDCSTPCSGSWLRDVWDWICT----VLTDFKTWLQSKLLPR 2002

Qy   1897 IPGCPFYSCQOKGYKGPWIGSGMLQARCCPCGAELIFSVENGPAKLYKGPRTCSNYWRGAVP 1956
          :||| |.||| :||:| |.|.| :| .| ||| :.| || ||. :::||||||| |:|:.|
Db   2003 LPGDPFFSCQORGYRGVWRGDGVMQTTCPCGAQITGHVKNGSMRIV-GPKTCSNTWHGTFP 2061

Qy   1957 VNARLCGSARDPT-DWTSLVVNYGRDVDCYKEYKQMGDHIFVTAVSSPNV-CFTQVPPTLR 2014
          :||  .  ||.|  :   :.. | |:  |  .  :.|  ||| |:::||  |    |.| 
Db   2062 INAYTTGPCTPSPAPNYSRALWRVAAEEYVEVTRVGDFHVVGTMTDNVKCPCQVPAP-E 2120

Qy   2015 AAVAVDGVQVCY--------------LGEPKTPWTTSACCY-GPDGKGKTVKL--PFR 2056
          . . .||: : |               .|.  . .:.. ||  |:   | :|:  ||
Db   2121 FFTEVDGVRLHRYAPACKPLLRDEVTFQVGLNQFPVGSQLPCEPEPDVTVLTSMLTDPSH 2180

Qy   2057 VDGHTPGVRMQLNLRDALETND-----------CNSTNNTPSDEAAVSALVFKQEL--- 2101
           :|:||||:   :| |  |              |.||     :|:|.|.:.|||:   
Db   2181 ITAETAKRRLARGSPPSLASSSASQLSAPSLKATCTTRHDSPDADLIEANLLWRQEMGGN 2240

Qy   2102 ---RRTNQLL-----EAISAGVDTTKLPAPSIEEVVVRKRQFRARTGSLTLP--PP--- 2147
              || ::|      | :| ||:|.|   |:.||   ::.:|  ::   | |  ||
Db   2241 ITRVESENKVVILDSFEPLRAEEDEREVSVAA--EILRKSRKFPPALPIWARPSYNPPLL 2298

Qy   2148 -----PRSVPGV--SCPESLQRSDPLEGPSNLPPSPP--------VLQLAMPMPLLGAGEC 2193
                |. .     .|  . .::.|.:| ||||||        | |.| |  |::|:. |
Db   2299 ESWKDPDYVPPVVHGCPLP--------PTMAPPIPPPRRKRTVVLTESTVSSALAELAT 2349

Qy   2194 NPFTAIGCAMTETG--GGPDDLP---------SYPP----KKEVSEWSDESWSTAT 2234
          :  : .| |  :|   |||| || :       |.||     ||:|:||| ||||:|
Db   2350 KTFGSGSGSAAVDSGTATAPPDQPSDDGDRGSDDESYSSMPPLEGEPGDPDLSDGSWSTVS 2409
```

```
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepacivirin #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2240,23

Query Match      20.3%;   Score 3126.5;  DB 1;  Length 3011;
Best Local Similarity  29.2%;  Pred. No. 5.1e-186;
Matches  951;  Conservative  434;  Mismatches 1219;  Indels  651;  Gaps   109;
```

```
Qy   2674  NPSAAWIGYLIHYPCLWSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSI  2733
                :||   ||| |   |      ||        ::|
Db   2823  TPVNSWLGNIIMFAPTLWARMLMTHFFSVLIARDQEQALNCEIYGACYSIEPLDLPPI  2882

Qy   2734  IAGVHGIEAPSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLA  2793
                ||   |||| |   |   ||||   |
Db   2883  IQRLHGLSAFSLHSYSPGEINRVAACLRKLGVPPLRAWRHRAWSVRARLLARGGKAAICG  2942

Qy   2794  RPLL-WHATSR---PLPDLDKTSVARYTTFNYCDVVSPEGDVFITPQRRLQKFLVKYLA  2848
                || |            |:|           :::           | |
Db   2943  KYLFNWAVRTKLKLTPITAAGRLDLSGWFTAGYSG-----GDIYHSVSHARPRW------  2991

Qy   2849  VIVFALGLIAVGLAI  2863
                |  |||  || |
Db   2992  -FWFCLLLLAAGVGI  3005
```

RESULT 11

JC5620
genome polyprotein - hepatitis C virus (isolate EUH1480)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu:
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: JC5620
R;Chamberlain, R. W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem. Biophys. Res. Commun. 236, 44-49, 1997
A;Title: The complete coding sequence of hepatitis C virus genotype 5a, the predominant g
A;Reference number: JC5620; MUID:97366593; PMID:9223423
A;Accession: JC5620
A;Molecule type: mRNA
A;Residues: 1-3014 <CHA>
A;Cross-references: UNIPROT:O39928; GB:Y13184
A;Experimental source: genotype 5a, which predominates in South Africa
A;Note: the translation of the nucleotide sequence is not complete in this paper
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <MEE>
F;192-389/Product: major envelope protein E #status predicted <EPM>
F;384-408/Region: hypervariable #status predicted
F;390-730/Product: nonstructural protein NS1 #status predicted <NS1>
F;731-1007/Product: nonstructural protein NS2 #status predicted <NS2>
F;1008-1616/Product: hepacivirin #status predicted <NS3>
F;1231-1238/Region: nucleotide-binding motif A (P-loop)
F;1313-1318/Region: nucleotide-binding motif B
F;1317-1320/Region: DEXH motif
F;1617-1863/Product: nonstructural protein NS4a #status predicted <N4A>
F;1864-2014/Product: nonstructural protein NS4b #status predicted <N4B>
F;2015-3014/Product: nonstructural protein NS5 #status predicted <NS5>
F;2210-2249/Region: interferon sensitivity determining #status predicted

Query Match   20.3%; Score 3119; DB 1; Length 3014;
Best Local Similarity 29.4%; Pred. No. 1.5e-185;
Matches 945; Conservative 448; Mismatches 1251; Indels 574; Gaps 105;

```
Qy     8  TSPVPAPRTRKNKQTQASYPVSIK------------------------TSVERGQRAKRKVQORDARPR   51
             |:||  |  |  |  :                                  ::  :| |
Db     3  TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPKLGVRATRKNSERSQPR   59

Qy    52  NYKIAGIHDGLQTLAQALPA----------------------------HGWGROD   79
             ||       |   |                                 || ||
Db    60  GRR-------QPIPKARRPTGRSWGQPGYPWPLYANEGLGWAGWLLSPRSSRPNWGPND  111

Qy    80  PRHKSRNLG------ILLDYP--LGWIGDVTTHTPLVGPLVAGAVVRPVCQIVRLLEDGVN  132
             ||:| ||             | |   |    |       |    |
Db   112  PRRKSPNLGRVIHTLTCGFPHLMGYI--------PLVGGPV-GGVSRALAHGVKVLEDGIN  163

Qy   133  WATG---WFGVHLFVVCLL-SLACPCSGARVTDPDTNTT---ILTNCCQRNQVIYCSPST  185
             |         | || |||   :     |   :    :   :::        :  ::
Db   164  YATGNLPGCPFSIFVLALLWCLTVPASAV----PYRNASGVYHVTNDCPNSSIVYEADNL  219
```

```
Db   1773  QYLAGLSTLPGNPAIASLMAFTAAVTSPLITGQTLLFNILGGWVAAQLAAPGAATAFVGA  1832
                ||      |    | : |            :|::|
Qy   1726  GAAGTALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNP  1782
                |   || |  |       |          |  |   | |    |  || :: :|
Db   1833  GLAGAAL---DSVGLGKVLVDILAGYGAGVAGALVAPKIMSGEVPSTEDLVNLLPAILSP  1889

Qy   1783  AAGVVCGVLSACAMFALTTAGPD----HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGI  1838
                | |||| |  |     :                 |  |    :   |    ||  :
Db   1890  GALAVGVVFASLL--RRRVGPEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAARVTAI  1947

Qy   1839  LEASTPWSVISACIRWLHPTEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIP  1898
                |   |  |   | | |  |  ||    |  |      |      |        || |
Db   1948  LSSLTVTQLLRRLHQWISSECTTPCS-GSWLRDIWDWICEVLSDFKTWLKA--KLMPQLP  2004

Qy   1899  GCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVN  1958
                |  | |  || |   |:          |       |  :   | |    ||    |:
Db   2005  GIPFVSCQRGYRGVWRGDGIMHTRCHCGAEITGHVKNGTMRIV-GPRTCKNMWSGTFFIN  2063

Qy   1959  ARLCGSARPDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFTQVP------  2010
                |  ||  ||                    :|           ||
Db   2064  AYTTGPCTPLPAPNYKFALWRVSAEEYVEIRRVGDFHYVSGMTTDNLKCPCQIPSPEFFT  2123

Qy   2011  ------------PTLRAAVD----GVQVQCYLGEPKTPWTTSACCYGPDGK  2047
                            ||  ||      | ||     |      |      ||
Db   2124  ELDGVRLHRFAPPCKPLLREEVSFRVGLHEYPVGSQLPC---EPE-------PDVA  2169

Qy   2048  GKTVKL--PFRVDGHTPGVRMQLNLRDALETND----------CNSTNNTPSDERAAVS  2093
                || :    | :|    | ||   | |  :              :: | |     || |
Db   2170  VLTSMLTDPSHITAEAAGRRLARGSPPSMASSSASQLSAPSLKATCTANHDSPDAELIEA  2229

Qy   2094  ALVFKQEL------RRTNQLL------EAISAGVDTTKLPAPSIEEVVRKRQFRARTGS  2141
                |:::|          ||             :      |       | :  | | | |
Db   2230  NLLWRQEMGGNITRVESENKVVILDSFDPLVAEEDEREVSVPA--EILRKSRRFAPALPV  2287

Qy   2142  LTLP---------PPPRSVP----------------------  2152
                ||||                                       
Db   2288  WARPDYNPLLVETWKKPDYEPPVVHGCPLPPPRSPPVPPPRRKKRTVVLTESTLPTALAEL  2347

Qy   2153  --------GVSCPESLQRSDPLEGSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGC  2201
                          || |      |            :|  |       |
Db   2348  ATKSFGSSSTSGITGDNTTTSSEP--APSGCPPDSDVESYSSMPPLEG---------  2393

Qy   2202  AMTETGGGPDDLPSYPKKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKR  2261
                           |             :                  |
Db   2394  ----EPGDPD-----------LSDGSWSTVSS--------GADT-----------  2414

Qy   2262  PTKKKLGKSEFSCSMSYTWTDVISFKTASKVLS-ATRAITSGFLKQRSLVVTEPRDAEL  2320
                |                   :   | |     |      |      |
Db   2415  -------EDVVCCSMSYSWTGALVTPCAAEEQKLPINALSNSLLRHHNLVYSTTSRSACQ  2467

Qy   2321  RKQKVTINRQPLFPPSVHKQVRLAKEKAKSKVVGVMWDYDEVAAHTPSKSAKSHITGLRGT  2380
                |  |  |     |   |   | |      | |           |  |  |    |  |
Db   2468  RKKKVTFDRLQVLDSHYQDVLKEVKAAASKVKANLLSVEEACSLAPPHSAKSKF-GYGAK  2526

Qy   2381  DVRSGAARKAVLDL----QKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISY  2436
                |   ||   |        :|      |      | |    | ||      |    | :|
Db   2527  DVRC-HARKAVAHINSWKDLLEDSVTPI---DTTIMAKNEVFCVQPEKGGRKPARLIVF  2582

Qy   2437  PHLEMRCVEKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMW--SPDAVGATCDTV  2493
                |          |          |  |        | || |:       |       |
Db   2583  PDLGVRVCEKMALYDVVSKLPLAVMGSSYGFQYSPGQRVEFLVQAWKSKKTPMGLSYDTR  2642

Qy   2494  CFDSTITPEDIMVETDLYSAAKLSDQHRAGITHIARQLYAGGPMIAYDGREIGYRRCRSS  2553
                ||| :||:|   | | |    |  |   | |    | |  || |   || |  |  | |
Db   2643  CFDSTVTESDIRTEEATYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRAS  2702

Qy   2554  GVYTTSSSNSLTCWLKVNAAAEQAGMKNPREFLICGDDCTVIWKSAGADADKQAMRVFASW  2613
                ||    |::|        | |                |   |   |     |
Db   2703  RVLTTSCGNTLTRYIKARAACRAAGLQDCTMLVCGDDLVVICESAGVQEDAASLRAFTEA  2762

Qy   2614  MKVMGAPQDCVPQPKYSLEELTSCSSNVTSGITKSGKPYIYLTRDRIPLGRCSAEGLGY  2673
                |::    |  | |   :   |  || | |   |  | | :| ||        |    |
Db   2763  MTRYSAPPGDPPQPEYDLELIITSCSSNVSVAHDGAGKRVYYLTRDPTTPLARAAWETARH  2822
```

Left panel:

```
Qy   186 CLHEPGCVICADE------CWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240
Db   220 ILHAPGCVPCVLEDNVSRCWVQITPTLSAPS-FGAVTALLRRAVDYLAGGAAFCSALYVG  278

Qy   241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGW---MAGKVERVIF  297
Db   279 DACGALSLVGQMFTYKPRQHTTVQD---CNCSIYSGHITGHR--MAWDMMMKMWSPTTALL 333

Qy   298 LTKLASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLALMLY------IEATS      348
Db   334 MAQLL-RIPQVVIDIIAGGHWGVLLAAAYFASTANWAKVILVLFLFAGVDGRTHTVGGTV  392

Qy   349 GNPIRVPT------------------------------------------GC          358
Db   393 GQGLKSLTSFFNPGPQRQLQFVNTNGSWHINSTALNCNDSLQTGFIAGLMYAHKFNSSGC  452

Qy   359 -SIAEFCSPLM1---------------------PCPCHSYLSENV-SEVICYSP         389
Db   453 PERMSSCRPLAAFDQWGWGTISYATISGPSDDKPYCWHYPPRPCGVVPARDVCGPVYCFTP 512

Qy   390 KWTRPITLEYNN-----------SISWYPYTIPGARGCMVKFKNN----TWGCCRIRNVPSY-- 436
Db   513 S---PVVVGTTDRRGCPTYNW-----GSNETDILLLNNIRPPAGNWFGCTWMNSTGFVK    563

Qy   437 ------CTMG-----------------TDAVWNDTRNTYEACGVTPWLT------TAWHNGSAL 471
Db   564 NCGAPPCNLGPTGNNSLKCPTDCFRKHPDATYTRCGSGPWLTPRCLVHYPYRLWHYPCTV  623

Qy   472 KLAILQ---YPGSKE--MFKPHNWMSG-HLYPEGSD-----TPIVYFYDPVNSTLLPPERW 521
Db   624 NYTIFKVRMFIGGLEHRLEAACNWTYGERCDLEDRDRAELSPLH----------TTTQW   673

Qy   522 ARLPGTPPVVRGSWLQVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTK  581
Db   674 AILPCS-----------FTPTPALSTGLIHLHQNIVDTQYLYGLSSSIVSWAVKWE      718

Qy   582 AVVLILLGLCGSKYLILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPF  641
Db   719 YIMLVFL-----------------------------------------LLADARICTC    735

Qy   642 ALIFFICCYLRCRLRYAALLGFVPMAAGLPLTFFVAAAAQPDYDWMVRLLVAGLVLM---  699
Db   736 LLILLLICQAEATCKNVIVLN--------------------ARAAAGNHGFFWNGLLVVC--LAWHV 779

Qy   700 AGRNRGHRIALLVGPWPLVALLTLLHLVTPASAFDTELIGGLTIPPVVALVMSRFGFFA   759
Db   780 KGRLVPGATYLCLGVWPLL----LVRLLRPHRALALDSSDGGTVGCLV-LIVLTIFTLTP  834

Qy   760 HLLPRCALVNSYLWQRWENWFWNVTLRPERFFLVLV---CFPGATYDALVTFCVCHVAL-  815
Db   835 GYKKKVVLV---MW----WLQYFIARVEAIIHVWVPPLQVKGGRDAVIMLTCLFHPALG   886

Qy   816 --------------LCLTSSAASFFGTDSRVRAHRMLVRLGKCHAWYSHYVLKFFLLVFGE 862
Db   887 FEITKILFGILGPLYLLQHSLTKVPYFLRARALLRLCLLAK-HLVYGKYVQAALLHLGRL  945

Qy   863 NGVFFYKHL--HGDVLPNDFASKLPLQEPFFFPEGKARVYRNEGRRLACGDTVDGLPVVA  920
Db   946 TGTYIYDHLAPMKDWAASGLRELTVATEPIVFSAMETKVITWGADTAACGNILAVLPVSA 1005

Qy   921 RLGDLVFAGLA--MPPDGWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGS  978
Db  1006 RRGREIFLGPADDIKTSGWRLLAPITAYAQQTRGVLGAIVLSLTGRDKNEAEGEVQFLST 1065

Qy   979 LATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCS 1038
Db  1066 ATQTFLGICINGVMWTLFHGAGSKTLAGPKGPVQMYTNVDKDLVGNPSPPGKGSLTRCT  1125

Qy  1039 CGETKGYLVTRLGSLVENKSDDPYWCVCGGALPMAVAKGSSGAPILCSSGHVIGMFTAA- 1097
Db  1126 CGSADLYLVTRHADVIPARRRGDTRASLLSPRPISYLKGSSCGGPIMCPSGHVGVVFRAAV 1185

Qy  1098 --RNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILIAPTGSGKSTKLP 1155
```

Right panel:

```
Qy  1186 CTRGVAKALEFVPVENLETTMRSPVFTDNS--------TPPAVPHEFQVGHLHAPTSGSGKSTKVP 1242
Db  1156 LSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYLT 1215

Qy  1243 AAYAAQGYKVLVLNPSVAATFGFGAYMSRAYGVDPNIRTGVRTVTTGAGITYSTYGKFFA 1302
Db  1216 -GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHAN 1273

Qy  1303 DGGCSGGAYDVIICDECHSQDATTILGIGTVLDQAETAGARLVVLATAIPPGSVTTPHPN 1362
Db  1274 ITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELARKGITAVSYYRGC 1333

Qy  1363 IEEVALPSEGEIPFYGRAIPLVLIKGGRHLIFANQKKKAKETAKGMNKPGRKAVAYYRGL 1422
Db  1334 DISKIP-EGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSA 1392

Qy  1423 DVAVIPATGDVVVCSTDALMTGFTGDFDSVIDCNSAVTQTVDFSLDPTFTIETTTVPQDA 1482
Db  1393 IVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTILDT 1452

Qy  1483 VSRSQRRGRTGRGRHGIYRYVSSGERPSGIFDSVVLCECYDAGCAWYDLTPAETTVRLRA 1542
Db  1453 YRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNVYLLTAAQLQLCHQYGVAA 1511

Qy  1543 YLNTPGLPVCQEHLEFWEGVFTGLTNIDAHMLSQAKQGGENFPYLVAYQATVCVRAKAPP 1602
Db  1512 PNDAPRWQGARLGKK-------PCGVLWRLDGADACPGPEPSE-------VTRYQM-CFT--E 1557

Qy  1603 PSWDTMWKCMICLKPTLTGPTPLLYRL-------GAVQNEITLTHPITKYIMACMSADLE 1655
Db  1558 VNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI------- 1611

Qy  1656 VITS-TWLVGGVVAALAAYCL-TVGSVAI----VGRIILSGRPAITPDREVLYQQFDE   1708
Db  1612 VEECASFIPLEAMVAAI-------DKLKSTITT--SPFTLETAL-------EKLNTFLGPHAAT 1660

Qy  1709 MEECSAGSLPYVDEARAIAGQFKEKVLGLGTAGQKAETLKPAATSMWSKAEQFWAKHMWN 1768
Db  1661 ILAIIEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGAL 1720

Qy  1769 FVSGIQYLAGLSTLPGNPAVATLMSFTAAVTSPLTTHQTLLFNILGGWVASQIAPPTAAT 1828
Db  1721 AFMMAGRAGTALGTWTSVGFVFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVVYSAF 1780

Qy  1829 AFVVSGNMAGAAVNIGLGRVLIDILAGYGTGVAGALVAFKIMCGERPTAEELVNLLPSIL 1888
Db  1781 NPAAGVGVLSACAMFALTTAGPD---HWPNRLLTMLARSNTVCNEYFIATRDIRRKIL   1836

Qy  1889 CPGALVVGVI--CAAVLRRHIGPGEGAVQWMNRLIIAFASRGNHGSPTHYVPETDASAKVT 1946
Db  1837 GILEASTPWSVISACIRWL----HTPTEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQ 1892

Qy  1947 QLLSLSLIVTSLLKRLHTWIGEDYSTPCDG----TWLRAIWDWVCT---ALTDFKAWLQ  1997
Db  1893 S--MVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNY 1950

Qy  1998 AKLLPQLPGVPFFSCQKGYKGVWRGDGVNSTKCPCGATISGHVKNGTMRIV-GPKLCSNT 2056
Db  1951 WRGAVPVNARLCGSARPDT--DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFTQ 2008

Qy  2057 WQGTFPINATTTGPSVPAPNYKFALWRVGAADYAEVRRVGDYHYITGVTQDNLKCPCQ   2116
Db  2009 VPPTLRAAVAVDGVQVQCYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGVRMQL 2068

Qy  2117 V-PSPEFFTELDGVRIHRF-APPCNPLLRE-----------EVTFSVGLHSYVVGSQL   2161
Db  2069 NLRDALETNDCNSTNNTPSDEAAVSALVFKQELRRTNQLLEAISAGVDTTKLPAPSIE-- 2126

Qy  2162 PCB---PEPDVTVLTSMLSDPAHITAETAK---RRLNRGSPPSLANSSASQLSAPSLKAT 2215
Db  2127 ---------EVVVRKRQFR-ARTGSLLTPPPPRSVPGVSCPESLQRSDP---------- 2165
```

```
Query Match          13.4%;     Score 2055;  DB 2;    Length 3005;
Best Local Similarity  25.7%;   Pred. No. 4.1e-119;
Matches 820; Conservative 401; Mismatches 1237;  Indels 736;  Gaps 122;

Qy   169 LTNCCQRNQVIYCSPSTCLHEPGCVICADECWVPANPYIS----HPSN------WTGTD  217
Db    33 LTNCCSDSEITACFDGGCLVSVGCTVCDGRCWDLYRPGVATRPGHPGGELFGALMSTSSP  92
Qy   218 SFLADHIDFVMGALVTCDALDIGELCGACVLVG-----DWLVRHWLIHIDLNETGTCYLE  272
Db    93 AVVAAYVAGITG-------LGEPFSAAALLAALARVEWTPR-----PNVTCLMD  134
Qy   273 VPTGIDPGFLGFIGW-MAGKVEAVIFLTKLASQVPYA--IATMFSSVHYLAVGALIYYAS  329
Db   135 CDLALQSE-----AWDLVEDMRDTFWTVKWVMTLPWRLWVGLVGMSFALMVVVLLLEQ  189
Qy   330 RGKWYQLLLALMLYIEA---------TSGNPIRVP------TGCSIAEFCSPLMIPC  371
Db   190 RVVMVLIVLAMAGSVBPFTGGCRCHVNGSIVAIPGAKPEDRPNATQVCVCPFGQMYWLPA  249
Qy   372 PCH--SYLSENVS----------EVICY--SPKWTRPITLEYNNSISWYPYT  409
Db   250 LCAGLAWRNGDVSGTTRDLPVRCPESMHGRGSVMCVWGSAHWSWRLATEVR---LW--EQ  304
Qy   410 IPGARGC---------------MVKFKNNTWGCCRIRNVPSYCTMGTDAVNDTRN--  450
Db   305 LPGSALCHFFAVGTTDRPHPATDVLSTHGIPCASCVVDRRASWCGNCTRDCWEKTGNKRL  364
Qy   451 TYEACGVTPWLTT-AW---HNGSALKLAI------LQYPGSKEMFKPHNWMSGHLYFEGS  500
Db   365 SFEACGLGSRLTAELWAHVVDGGTESKIVTDVGERPRYPSSHGVGTHHTAVVAARNTVS  424
Qy   501 D---------------TPIVYFYDPVNSTLLPPERWARL----PG---TPPVVRG  533
Db   425 DIGGYWHAIACPNPPSPSALPKLIPGRPVNACLTICEKGRLNTAWQAPGGXFAPIFTEC  484
Qy   534 SW------LQVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTK--AVVLI  586
Db   485 NWPRRSCVRVCEGYAPFDFPGSKTGFI---RSRAGFQQIARAGSMPHPRWLLTDYLFILLV  541
Qy   587 LLGLCGSKYLIIAYLCYLSLCFGRASGYPLR---PVLPSQSYLQAGWDVL------SK  635
Db   542 LMKLLAEARLVPLVAVALYWWFNNMAEAATLRILHPVNITTAAPMTWPALPVPTVVCPTK  601
Qy   636 AQVAPFALIFFICCYLRCRLRYAALLGFVPMAAGLPLTFFVAAAAAQPDYDWW--VRLLV  693
Db   602 AGAVGDAL--------YMAACLAGNATLGVVSSVWG-GAYGAEAGARGLWRWWGGFRSFC  653
Qy   694 AG-----------------LVLWAGRNR---------GHRI  708
Db   654 AGLAWLTNVGAYLPVVEAAVGPEFVSAPLLVWAWBDSIAVFILMAANIWAYWSDHLGVKL  713
Qy   709 ALLV-----GPWPLVALLTLLHLVTPASAFDTEIIGGL---------TIPPVVALV  750
Db   714 AGLVAAHLAKGALPLVLLVAAXVTRHRHSVLGFEVCVSLDADASPSDWSWWALAAAVSWC  773
Qy   751 VMS-----------RGFFA-----HLLPRCALVNSYLWQRWENWFWNVTLRPERPFLVL  794
Db   774 LLTLGLTHGGRLRKLEFYSYWCRFHQAVRMRLLCSKVGRGGRDW----------RVTAVW  824
Qy   795 VCFPGATYDALVTFCVCHVALLCLTSSAASFPGTD----------SRVRAHRM--LVR  840
Db   825 VA-AGIIFPREVVRCSA----ALTALAALLDSIDYLLETLILTLTAAQPARARLLDSLTF  878
Qy   841 LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQOBPFFPPEGKAR--  898
Db   879 LGDAD------LTRAFPVRRLERRGVTLFQHC-GQVSXGAAAILXDLGVALEPVSVTARDC  931
Qy   899 -VYRNEGRRLACGDTVDGLPVVARLGDLVFAGL----AMPPDGWAITAPFTLQCLSERG  952
Db   932 YIVRDAARTLACGQRVEGLPVVARRGDEVLVGVFPSVRALPP-GFVPTAPVVVM--QRG  987
Qy   953 T--LSAMAVVMTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGS 1010

Db  2216 CTIQGHHPDADLIKANLLWRQCMGGNITRVEAAENKVEILDCFKPLKEEBDDREISVSADC 2275
Qy  2166 -LEGPSNLPPSPPVLQLAMP---MPLLGA--GECNPFTAIGCAMTETGGGPDDLPSYPP 2218
Db  2276 FKKGPA-FPPALPV--WARPGYDPPLLETWKRPDYDPPQVWGCPIPPAGPPVPLPR--R 2330
Qy  2219 KKEVSEWSDESWSTATTASSYVTGPPY----PKIRGKDST------QSAPAKR------ 2261
Db  2331 KRKPMELSD---STVSQVMADLADARFKVDTPSIEQDSALGTSSQHDSGPEEKRDDNSD 2387
Qy  2262 -----------PTKKKLGKSEFS--------CSMSYTWTD-VISFKTASKVLSATR 2297
Db  2388 AASYSSMPPLEGEPGDPDLSSGSWSTVSGEDNVVCCSMSYTWTGALITPCSAEEEKLPIN 2447
Qy  2298 AITSGFLKQRSLVYYTEPRDAELRKQKVTINRQPLFPPSYHKQVRLAKEKASKVVGVMWD 2357
Db  2448 PLSNTILRHHNLVYSTSSSRAGLRQKKVTFDRLQVLDDHYREVVDEMKRLASKVKARLLP 2507
Qy  2358 YDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDL---QKCVEAGEIPSHYRQTVIV 2413
Db  2508 LEEACGLTPPHSARSKY-GYGAKEVRS-LDKKALKHIEGVWQDLLDDSDTP---LPTTIM 2562
Qy  2414 PKEEVFVKTQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAVMGDAYGF-VDPRT 2472
Db  2563 AKNEVFAVEFSKGGKGKKPARLIVYPDLGVRVCEKRALYDVAQKLPTALMGPSYGFQYSPAQ 2622
Qy  2473 RVKRLLSMWSPDAV--GATCDTVCFDSTITPEDIMVETDIYSAAKLSDQHRAGIHTIARQ 2530
Db  2623 RVDFLLKAWKSKKIPMAFSYDTRCFDSTITBHDIMTEEESIYQSCDLQPEARVAIRSLTQR 2682
Qy  2531 LYAGGPMIAYDGREIGYRRCRSSGVVTTSSSNSLTCWLKVNAAAEQAGMKNPRFLICGDD 2590
Db  2683 LYCGGPMYNSKGQQCGYRRCRASGVFTTSMGNTMTCYIKALASCRAAKLRDCTLLVCGDD 2742
Qy  2591 CTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSNVTSGITKSGK 2650
Db  2743 LVAICESQGTHEDEASLRAFTEAMTRYSAPPGDPVPAYDLELVTSCSSNVSVARDASGN 2802
Qy  2651 PYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHFMEQMLFEDKL 2710
Db  2803 RIYYLTRDPQVPLAKAAWETAKHSPVNSWLGNIIMYAPTLMARIVLMTHFFSVLQSQEQL 2862
Qy  2711 PETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSLTDMTMPPLRA 2770
Db  2863 EKTLAFEMYGGSVYSVTPLDLPAIIQRLHGLSAFSLHSYSPSEINRVASCLRKLGVPPLRA 2922
Qy  2771 WRKKARAVLASAKRRGGAHAKLARFLL-WHATSR----PLPDLDKTSVARYTTFNYCDVY 2825
Db  2923 WRHRARAVRAKLIAQGGRAAICGIYLFNWAVKTKRKLITPLADADRLDLSSWFT------VG 2977
Qy  2826 SPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAI 2863
Db  2978 AGGGDIYHSMSRARPRNL------LLCLLLSVGVGI 3008
```

RESULT 12
T08841
polyprotein - douroucouli hepatitis GB virus A
C;Species: douroucouli hepatitis GB virus A
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
C;Accession: T08841
R;Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K.
J. Gen. Virol. 79, 41-56, 1998
A;Title: Genomic analysis of two GB virus A variants isolated from captive monkeys.
A;Reference number: Z16486; MUID:98120818; PMID:9460920
A;Accession: T08841
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1005 <RNA>
A;Residues: 1-1005 <ERK>
A;Cross-references: EMBL:AF023425; NID:g2828599; PIDN:AAC40502.1; PID:g2828600
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: polyprotein

```
Qy   988  LGFFSVVKTSMLGRDEREHEGSIVVLGTSTTRSMGTCVNGVMYTTFHGSNARTLAGPVGP 1047
              |  :||      ||  :  |   :|||     |  ||   ||:    ||  ||| ||| ||:
Db  1011  IHPITVDAANDQDIYQPPCGAGSLTRCSCGETKGYLVT------RLGSLVEVNKSDDP   1062

Qy  1048  VNCRWWNSPSDDVAVYPLPSGASCLEPCKCGTQSVWCIRNDGALCHGRLSKLVELD----  1102
              ::  |   |  :  ||| ||:||   | |    ||  |   | | ::  |    :|
Db  1063  YWCVCGALPMAVA--KGSSGAPILCSSGHVIGMFTAARNSGGSVSQIR-VRPLVCAGYHP  1119

Qy  1103  ------LPTEISDFRGSSGSPILCDEGHVVGMMVSVLHRGVKVTGVRYVKPWETLPKDS   1155
                     |:  |  ||||| |    ||:||  :  ||: |: |   ||||
Db  1120  QYTAHATLDTKPTVPNE--YSVQILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTAS  1177

Qy  1156  QVKSEA-----PPVPGKTGFTEAPLYLPTGSGKSTRIPLEYTKAGHKVLVLNPSIATVRA  1210
              |:  |        |||| |  ||| |||||||||:|||:||||:|| |||||| ||:|:|
Db  1178  MPKYMHATYGVNPNCYFNGKCT----NTGASLTYSTGVMYLTGACSRNY----DVIICDE  1229

Qy  1211  MGPYMEKLSGQHPSIYCGHDTTAYSRTTGSPLTYCTYGRFM--ANPRRYLRGADIVICDE  1268
              || || : |:||| | | |||  :  ||:||||  | |  |  | |||||||||:||||||
Db  1230  CHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANIETEIQLTDEGTIPFHG  1289

Qy  1269  CHVTDPTSVLGMGRARLLARECGVRLLLFATATPPGPALAQHESIKEVPLGVDGEVAFYG  1328
              || || :|||:| |||| ||| ||||:||||||||| :| |||||:||||||||||:|||
Db  1290  KKIKEENLKKGRHLIFEATKKHGIDELANELARKGITAVSYRGCDISKIPEGDCVVATD   1349

Qy  1329  HKLPVERYRTGRHLLFCHSKVECNRLHAALSTAGCNAVVYRGNE-QEIPAGDVCVCATD   1387
              : |  |   ||||| | |   || ||:|| || |||| |||| |  ||| ||||| |||
Db  1350  ALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGI  1409

Qy  1388  ALSTGYTGGFSTVTDCGLMVEEVEVEVTLDPTITISVRTTPAPAELRAQRRGRCGRGSQGT  1447
              ||:|||||  | | ||:|||||||:|| || |:|:|  | ||||||| |||||| ||| ||
Db  1410  YYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTILDTYRTOPGLPAIGANLDEW  1469

Qy  1448  YYYAMTASAPAGTLRSGPLWAAAVEAGVAWYNLEPDMTADILRAYDACPYTAAITASVGEA  1507
              ||| : | |   | ||| ||| |:|||:||| | ||:|||||||||||||| ||:|| |||
Db  1470  ADLFSMVNPEPSF--VNTAKRTADNYVLLTRAAQLQLCHQYGYAAPNDAPRWQG-ARLGKK  1526

Qy  1508  INFFSGLVPMRNYPQVAWAKSHQHNWPLLVGVQRTMCQEANVAGPGDGPEWAGLAGTGPI  1567
              |  |  || |     | |  |  ||| ||| | || || |||| | | |||||| |||
Db  1527  PCGVLWRLDGADACPGPEPSE--------VTRYQMCFTEVNTSGTAALAVGVGVAMA    1575

Qy  1568  PLLCRW--GARPPPSVAPHHWVDDLQARLGVAEGYSPCY------AGPILLVGLALAGG   1618
              | | |    |:| | | | || ||||||| ||||:|||       |||:|||||| |||
Db  1576  YLAIDTFGATCVRRCWSI-----------TSVPTGATVAP--------VVDE        1608

Qy  1619  AVLAHWTGSLVVVTSWRVNGNGNPLIQQSTRGVSTSAPYQLTVCVEGEQTPADGKCAAEA  1678
              || | |||| | | || | || |||||  ||| ||:||||||| ||||:|| | ||||||
Db  1609  EBIVEECASFIPLEA-------MVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATIL  1662

Qy  1679  VQLXESTCGWGPMAAASFDCAGMKGVLDSMRTTAA--------AAVEKSDSLWRSFCAN--  1728
              | |:| | |||| |||  | |||| | ||:| |||        || |||  |||  | |
Db  1663  AIIBYCCGLVTLPDNPFASCVPAFIAGITTP--------LPHKIKMFLSLFGGAIASKLT  1714

Qy  1729  --NYC--------PPGGGATSASAFFASLDTKFAQAWDAIFTNGRSLLVGAGYGARRNP  1779
               :|         ||   | || |||||| |:||||||||||||||||| | ||||:||
Db  1715  DARGALAFMMA-----------------GAAGTALGTWTSVGF----------        1740

Qy  1780  PLGVAAAFLMGMSAGHQVHVRLAAALLLGVGGTMLGT-PSVGLAMSGAYPAGGSITSSWL  1838
              ||||||| || || ||||||||||||||||||||||| || |||||||| |||||:|||
Db  1741  --VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFAL  1798

Qy  1839  SAIVAVLGGWEGAXNAASLTFDFLTGRAELKD-MWFLVSCKASPGASVAGVALGLLLWSM  1897
              ||:|||||||| ||:||||||| |||||||| | || |||:||||| |||||||||||||
Db  1799  TTAGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKIIGILEASTPWSVISACIRWLHT-  1857

Qy  1858  -----PTEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKG  1911
                   | | | | | |||||||||||||||||||||||||||||||||||||||||||
Db  1898  KKKGVGEDWVNRILITLLPRGSVLPDGFFVKS-EFTERVSTILRK-------MSLSRWVMTL  1949

Qy  1912  PWIGSGMLQARCPCGAELIFSVENG-FAKLYKGPRTCSNYWRGAVPVNARLCGSARPDT  1970
              ||| | || |||||||||||||||| |||||||||||||||||||||||||||||||||
Db  1950  VERRELDLETPCSSMLWDLIDW-----LVRFGRYIGRRLKGMVPSVRVPLVGCTPGWGG  2003
```

```
Db  2004  SWVGEGHIEARCACGCIITADVEEGKLVDVHYSSRLCSNYLKGTVPVSAAGSGDAEPEVP  2063
              |: |: |     | :    | |||| |||:   |||:|||: |||||   ||| || |
Qy  1971  DWTSLVVNYGVRDYCKYEKMGDHIFTAVSSPNVCFTQV-----PPTLRAAVAVDGVQV    2024

Db  2064  AGPAL-YQIGVAEWVQLVRKDKTLVVGASSVYHLHIDELRRAIRGPPMFVGGV---GVSW  2119
              |||| |:||  : | || | |  :  |  |  :  || :||||  ||| || |   | |
Qy  2025  QCVLGEPKTPWTTSACCYGPDGKGKTVK---------LPPRVDGHT---------PGVRM-  2066

Db  2120  EAPLQQPPLVVRA---------GQSVRFDDVRYSLP-----HTLALPPPRPPPPLAMP   2164
                :  | || :|          | :|| | ||| |      | ||| | | ||| ||
Qy  2067  ------------QLNLRD----ALETNDCNSTNNTPSDEAAVSALVFKQELRRTNQLL   2108

Db  2165  PPPPPEQVWTEBEADLREARARAIEA--VNERLPVVNPEAAQAAL-----------   2209
              ||| |:| | |  || || |||  |   |  || | ||| |||||
Qy  2109  EAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPRSVPG--VSCPES-----LQ  2161

Db  2210  -------------DALEEAAV-----------SLLPHVGAILGDDCSCEESFGGHFIP  2243
                           | ||| |          |||| | |:| | | |  | | || |
Qy  2162  RSDPLEGPSNL-------PPSPPVLQLAMPMPLLGA------GECNPFTAIGCAMTETGGGP  2210

Db  2244  EPDAVDVPIGMEVQVGPLRDQARDLGDRLAVLGARLESLAEAHPEASLNTERRTMGELI  2303
              | ||| |||| ||||| ||| |||| | ||| ||||| || | ||| |||   |||| |
Qy  2211  DDLPSYPPKKEVSEWSDESWSTATTASSYVTGP-----------PYPKIRGKD----   2252

Db  2304  DTLADVQARLDVTCRSDTSGSSFEQISLSDSEPETIVEGGLKLEVVRPQP-VRFKDLIRP  2362
              | ||| |||| || | | || | |:| | | ||| |||| || |||| | |||| |||
Qy  2253  -----------STQSAPAKRPTKKKLG-------------------------       2268

Db  2363  GEGAKRLVTVRQSCCADRSATRAFALSLPIAAVTATLSFDLTDHTVSDSTGRVLDPLELL  2422
              |:| || | |:|||| || |||||||| | ||| |||||||||||| || || |||||||
Qy  2269  -------------KSEFSCSMSYTWTDVISFKTASKVLSATRAITSGFLKQRSLVYYTEPR  2316

Db  2423  QNAVGDLIVACRRSESSVSYSYIWSGAPLNTGRHQPAPMTRPIGTHITADTTKVYVTDPN  2482
              | ||| |:|||||||||:|||||||||||||||| | |||||||| ||| |||| | |||
Qy  2317  DAELRKQKVTINR-QPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHIT  2375

Db  2483  QAGERAAKVTIWRGSRVYDAHYRGVSVLQQAKTIKSPGWTYDEAIAKVHSRAAAGFGS   2542
              | || || | || | | || || | | ||| |:||| |:|||| |||| | |||  |
Qy  2376  GLRGTDVRSGAARKAVLD-LQKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPKKPPRLI  2434

Db  2543  KVTVGNMTTPAARAEVNNMLAKIKTKQEVP-----FTLVTKREVFF-----QKTTRKPRFI  2594
              || | | ||| || |||| |||||:||| |     |:|| ||| |      |||||| ||
Qy  2435  SYPHLEMRCVEKMYYGQVAPDVV-KAVMGDAYGF-VDPRTRVKRLLSMW-----SPDAVGA  2488

Db  2595  CFPPLDFRIAEKMILGD--PGLVAKGILGKSYLFQYTPNQRVKLMVDLWRQKKHPRAI--  2650
              | | || ||| | | |   ||| | | || || || ||:|| |:| ||:||||| ||
Qy  2489  TCDTVCFDSTITPEDIMVETDIYSAAKLSDQHRAGIHTIARQL---YAGGPMIAYDGREI  2545

Db  2651  TVDATCFDSSIDRDMAVETEVPAAASPNPD---------LVRALGSYYAEGPMVSPKGVPL  2703
              | |: ||||:| || ||||||||||:|| |         |||||||| ||||| | | ||
Qy  2546  GYRRCRSSGVYTTSSNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGAD-ADK  2604

Db  2704  GVRKCRSSGVLTTSSANSITCYIKVKAAARAVGLVDPDFLIAGDDCVIIYEDDGEDHAD-  2762
              | |:|||||:|||| ||||| | ||:||| ||| | | |||||||:|| | || |  |
Qy  2605  QAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCGSSNVTSGITKSGKPYYFLTRDPRIPLG  2664

Db  2763  -ALRV----ALGNYGYDCKPKTHASLDTAESCSSYLAE-CNVGTERVWLSTDMRKPLA  2815
               ||||    ||||| |||||||| |||| ||||||| |  ||| ||| |||| |||||
Qy  2665  RCSAEGLGY-NPSAAWIGYLIHHYPCLW---VSRVLAVHFMEQMLFEDKLP-ETVTFDWY  2719

Db  2816  RAASE---YSDPVSSALGTIL-MYP--WHPIVRWLLPHILIMAFRGGGTPDDLVVCEVQ  2869
              | ||    |||| || | ||| ||   | |:|| || ||:|||  | |||| |||||:|
Qy  2720  GKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVL  2779

Db  2870  GNHYSFPLRVLPEVLVSLHGPRCLRVTADSTKTKMEAGAALRDLGMHTLAFYRKRAGNVR  2929
              | :| | | || | || | | | | | || |||| |||| ||||| || ||| |||||
Qy  2780  ASAKRRGGAHAKLARFLLMHATSRPLEPDLDKTSVARYTTFNYCDVYSPEGDVFITPQRRL  2839

Db  2930  TRLLRGGKGWGRLARALLMHPGLKEHP---PSIKSIPGFKMATPYEHHETVWYSGEKPP  2985
              ||||||| ||||||||||||| |||      || | |||| ||||||||| || ||||
Qy  2840  QKFLVKYLAVIVFA                                             2853

Db  2986  WYWNDVKCLFGLVCA                                            2999
```

```
RESULT 13

T08839
polyprotein - marmoset hepatitis GB virus A
C;Species: marmoset hepatitis GB virus A
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 17-Nov-2000
C;Accession: T08839
R;Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K.
J. Gen. Virol. 79, 41-45, 1998
A;Title: Genomic analysis of two GB virus A variants isolated from captive monkeys.
A;Reference number: Z16486; MUID:98120818; PMID:9460920
A;Accession: T08839
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: genomic RNA
A;Residues: 1-2970 <ERK>
A;Cross-references: EMBL:AF023424; NID:g2828597; PIDN:AAC40501.1; PID:g2828598
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: polyprotein

Query Match             13.2%;   Score 2029.5;  DB 2;  Length 2970;
Best Local Similarity   25.5%;   Pred. No. 1.6e-117;
Matches   826;  Conservative 386;  Mismatches 1240;  Indels 785;  Gaps 133;
```

```
Qy  134  ATGWFGVHLFVVCLLSLACPCSG----------ARVTDPDTNTTILTNCCQRNQVIYCSPS  184
                                                : :
Db    2  ASLWF----FVLLL----PLGGGVLFKSKHRCRVGD----RFFLSNCCSRDEVVFCFGD   48

Qy  185  TCLHEPGCVICADECWVPANPYI-----SHPSNWTGTDSFLADHIDFVMGALVTCDALDI  239
                                        |:|||  |||:|| :|  |  :  :
Db   49  GCLVSYGCTVCTQSCWKLYRPGVATRPGSAPGELLGK---FGSYVGPLASAAYAAGVLGL  105

Qy  240  GE------LCG------------ACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGI----  277
               |||              |:   |   | | :|  || |||| :  ||
Db  106  GEPFSLALAGGTLTRLGRLPNLTCHLKCDFAFSYPGISVDFDWAFKKILELPSRLWCGL  165

Qy  278  -DPGFLG-FIGWMAGKVEAVIFLTKLASQVPYAIATMFSSVHYLAVGA----LIYYASRG  331
              |  ||  | | :|  ||   :|:|: |       || : |    |||
Db  166  SRAGVLALLVVMMLGLEQRLVLALLLLIVVGEAYCDVFEACHCAHWGGRPPRITPWVRG  225

Qy  332  K----------WYQLLLALMLYIEAT-SGNPIRVPTGC---------SIAEFCS------  365
                          |:|||  |  |: :||| ||| |:        ||| ||
Db  226  NGSVECTCPFGRMVWNNPVLCSGLIWKDGSREGNVYDLPPVCPSEVLGTVAVMCRWGSAYW  285

Qy  366  -----------PLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIP  411
                       | |  | |: : || :|   | ||  || |   |  : | | | |
Db  286  AWRLGDWVELHRELPRSALCVFHSFTQR---------MNMWPKDV---------DFLTP  326

Qy  412  GARGCMVKFKNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLT----TAWH  466
            | |||   :|    |  | | :  | |  |  ||| :  |  |      ||
Db  327  EGTPC----ASCVVDQRPAWCGSCVRDCWETGGPDFASCGIGTQVTEHLEAVLVH  377

Qy  467  NGSALKLAILQYPGGSKEMFKPHNWMSGHL-----------YFEGSDTPIVYFYD  509
                   |  ||          :   |  :         :|   | |||: |
Db  378  GG--VESVVTTPRGERPKYLGHHGQQAYYGAVSKLSTNYTVTQIGGYWHALKCPCVEFPS  435

Qy  510  -----------PVNSTLL----PPERWARLPG---TPPVVRGSWLQVP-----QGFYSDV  546
                      |  :||    || || |||    ||||||:||  ||      | | ||
Db  436  GDLPRRIPGIPVNACLSGAGPHPFTAWA--PGGWYAPVFTKCNWPTTPGVDVCRGFAYDF  493

Qy  547  ----------------KDLATGLITKDKAWKNYQVLYS-----------------A  569
                            |:|   |  | :|| |:
Db  494  PGDHNGFIHVGARQQXYSGVMRSSPAWLLSDMAFALLVLMKVAEARLVPLFMLACWWWL  553

Qy  570  TGALSLTGVTTKAVVLILLG-----------LCGSK-----YLILAYLCYLSLCFGR  610
             || |:| |  |  ||  |           || |       : || | | |   |
Db  554  NGALAATITIIRPTVNLTTGAPPPLWTPPTVPTPTCSNQSSVGEALLAALCLSGQTLAG  613

Qy  611  ASGYPLRPVLP-SQSYLQAGW-------------DVLS---KAQVAPPAL-----  643
          ||   |  |   :| || |             ||||   | |   | |
Db  614  AAGDVWRVGNEGAQAVVKAGWYGWAGLRSFWRGLXWLQRAGDYLPAVQAALAPEYVAAPI  673

Qy  644  ------------IFFICCYL-----RCRLRYAALLGFVPMAAGLPLTFFVAAAAAAQ  682
```

```
Db  1638  TKRGVPIGG--PPIIPPPPTGEGERFPGEVKVVAERATETLBTACGWGPCSASFAYMKAC  1695
Qy  1624  -----MVAAIDKL------KSTITTTSPF--------TLETALEKLNTFLGPHAATILA  1663
Db  1696  EAGKLAAAKMDALCQAWSSGMFVPPAASAEHMPSIIQSLDTAFTRAWDTVFTHGRSLLV  1755
Qy  1664  IIEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFM  1723
Db  1756  GFTAAYG---AKRNPPLGVGASFLLGMSAGHSVHVRLAAALLLGVGGTMLGQASTGLA--  1810
Qy  1724  MAG---AAGTALGTWTSVGFVFDMLGGYAAAASSTACLTFKCLMGEWPTMDQLAGLVYSAF  1780
Db  1811  MAGAYFAGGSITSSWLSV--IVALLGGEGAVNAASLTFDLLSGRABEAKDAWC-IISCLA  1867
Qy  1781  NPAAGVGVLSACAMFALTTAGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILE  1840
Db  1868  SPGASVAGAALGVLLNSMKKGVGQDWVNRLLTLLPRSSVIPDDYFVKDEFV-VKVSSILR  1926
Qy  1841  ASTPWSVISACIRW-----------LHTPTEDDCGLIAWGLEIWQYCNFFVICFNVLKA  1889
Db  1927  R-----MSLSRWVMTLVDKREAEMETPASE---------LLWDFV-DWLVRLGRMLSR  1969
Qy  1890  GVQSMVNIPGCPFPYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGPAK-LYKGPRTCS  1948
Db  1970  RLSTLFPRIKVPLLGCSTGWSGHWEGSGHLETRCLCGCIVTGEIVDGEVQDLHYSTMWCS  2029
Qy  1949  NYWRGAVPVN-ARLCGSARPDP--------------TDWTSLVVNYG---VRDYCKYEKMGD  1992
Db  2030  NYIRKTIPVGVVGGAGAAEPAPPPRGLKTYQFGYKDWVEALVMEGTMTIHATSCYETTAA  2089
Qy  1993  HIFVTAVSSPNVCFTQ-----VPPTL----RAAVAVDGVQ--VQCYLGEPKTPWT  2036
Db  2090  ELRRAARTRP--CFSQGVACSWEAPMSPPAMVYRTGQTITIDGVRKMLPCTLVLGARPPP  2147
Qy  2037  TSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALV  2096
Db  2148  AAAVAPPPP---------TGWTQADEEDLQRTKAAAIEAIGTGLGLPSPEAAXAAL--  2194
Qy  2097  FKQELRRTNQLLEAISA-----GVDTT---KLPAPSIEEVVVRKR-----  2133
Db  2195  ----------DALEEAAVALLPHVPVIMGDDCSCRDEEAFQGHFIPEPNVSEVPIEPSVPDI  2246
Qy  2134  -QFRARTGSLITLPPPPRSVPGVSCPESLQRSDPLEGPSNLPPSPPVLQ--LAMPMPLLGA  2190
Db  2247  DHLRGVATSLITARLNDLEAMVAARAESLEDAKALSRSLSSDSDMPSLESTSSSSFEVISL  2306
Qy  2191  GECNPFTAI-----------GCAMTETGGGPDDLPSYPPKKEV  2222
Db  2307  SESEPETVIENGFPMEVQKPPLPXSFGGAKRILRVRQCQCCCDRS-----TLKAFPLELKV  2361
Qy  2223  SEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSC-------  2274
Db  2362  GE------VLFVCGYDPEGHRVTDDNGDLDMNTPIADTLGDIILSCETREEAQM  2409
Qy  2275  SMSYTWTDVISFKTASKVL--SATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINR-QP  2331
Db  2410  SYSYINSG--APLGCGRVPPITRPIGTHLTHDTTKVYVTDDRAAERAEKVTIWRGQR  2467
Qy  2332  LFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAAR-KA  2390
Db  2468  VYDEHYASVVGGEALKKAAATTSPGWTYAMAISKVRKRAAAGYSKVTVATLETGWPHVRT  2527
Qy  2391  VLDLQKCVEAGEIPSHYRQTVIVPKEBVFVKTPQKFRKPPRLISYPHLEMRCVEKMYYG  2450
Db  2528  MIDKIRCHE--EVPFTF-----VTKREVFF----SKTTRKPRFIVPPLDFRIAEKMILG  2577
Qy  2451  QVAPDVV-KAVMGDAYGF-VDPRTRVKRLLSMWS-----PDAVGATCDTVCFDSTITPEDI  2504
Db  2578  D--PGIVAKAILGPAYLFQYTPNQRVKLLVSTWESKTHPAAI--TVDATCFDSSIDERDM  2633
Qy  2505  MVETDIYSAAKLSDQHRAGIHTIARQLYAGGMIAYDGREIGYRRCRSSGVTTSSSNSL  2564
Db  2634  EVEAEIFAAA--SDQPEL-VKALCR--YYAEGPMVTPDGVPIGRKCRSSGVLTTSSANSI  2689
```

```
Qy  2565  TCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCV  2624
Db  2690  TCYIKVKAACAKVGLKDPSFFIAGDDCLIIYEDDGADPCERLRLALGNY------GYRCE  2743
Qy  2625  PQPKYSLELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGY-NPSAAWIGYL  2683
Db  2744  PSKHASLDTAECCSAYLAE-CTAGGSRRWWLSTDMRKPLARAAAE---YSDPVGSALGTI  2799
Qy  2684  IIHYPCLM----VSRVLAVHFMEQMLFEDKLP-ETVTFDWYGKNYTVPVEDLPSIIAGVHG  2739
Db  2800  L-MYP--WHPIVRYVLIPHVLIMAFRGGGTPDELVMCQVQGNTYSFPLKILPRVLVSLHG  2856
Qy  2740  IEAFSVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWH  2799
Db  2857  PWCLQVTTDSTKTRMEAGNALRDLGMKSLSWHRRRAGNVRTRLLRGGKAWGHLARCLLWH  2916
Qy  2800  ATSRPLPDLDKTISVARYTTFNYCDVYSPEGDVFITPQRR--LQKFLVKLVAVIVFAL  2854
Db  2917  PGLKEHP----PVIALPNFRLVTPYEHHEEVLISLKERPPWVKWLLGFLMALTAAL  2969
```

RESULT 14

T01075

polyprotein - hepatitis G virus (fragment) (isolate Japanese patient IM68)
C;Species: hepatitis G virus, HGV
C;Variety: isolate Japanese patient IM68
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T01075
R;Abe, K.; Kaneko, T.S.
submitted to the EMBL Data Library, October 1997
A;Description: Partial genome of hepatitis G virus isolated from Japanese patient.
A;Reference number: Z14246
A;Accession: T01075
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: genomic RNA
A;Residues: 1-1435 <ABB>
A;Cross-references: UNIPROT:O36196; EMBL:AB008344; NID:d1170988; PIDN:BAA23103.1; PID:d1
A;Experimental source: isolate Japanese patient IM68

Query Match   6.7%; Score 1028.5; DB 2; Length 1435;
Best Local Similarity 26.3%; Pred. No. 1.6e-55;
Matches 392; Conservative 218; Mismatches 646; Indels 235; Gaps 58;

```
Qy  1453  YRTQPGLPAIGANLDEWADLFSMVNP---EPSFVNTAKRTADNYVLLTAAQLCHQYGY  1509
Db     1  YDDCPYTAAVAADIGEAAVFFSGLAPLRMHPD-VSWAKVRGVNWPLLVGVQRTMCRETLS  59
Qy  1510  AAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNTSGTA----  1564
Db    60  PGPSDDPQWAGLK-GPNPVPLLLRW-GNDL------PSKVAGHHIVDDLVRRLGVAEGYVR  112
Qy  1565  -----ALAVGVGVAMAYLAIDTFGATCVRRCWSI----------TSVPTGATVAPVVD  1607
Db   113  CDAGPILMVGLAIAGGMIYASYTGSLVVVTDWDVKGGGNPLYRNGDQATPQPVVQVPPVD  172
Qy  1608  ----------EEBIVEBCASFIPL----------------EAMVAAIDKLKST  1634
Db   173  HRPGGESSSPSDAKTVTDAVAAIQVNCDWSVMTLSIGEVLALAQAKTAEAYTATAKWLAGC  232
Qy  1635  ITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFASCVFAFIAGITTPL  1694
Db   233  YTGTRAVPTVSIVDKL--FAGGWA----AVVGHCHSVI------AAAVAAYGASRSPPL  279
Qy  1695  PHKIKMFLSL-FGGAIASKLTDARGALAFMAGAAGTALGT----WTSVG-------  1739
Db   280  AAAASYLMGLGVGGNAQARLASA----------LLLGAAGTALGTPVVGLTMAGAFMGGASVSP  333
Qy  1740  ----FVFDMLGGYAAAASSTACLTFKCLMGEWPTMDQLAGLVYS---AFNPAAGVGVLSA  1792
Db   334  SLVTVLLGAVGGWEGVVNAASLVFDFMAGKLSTED-----LWYAIPVLTSPGAGLAGIALG  389
Qy  1793  CAMFALTTAGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACI  1852
```

S68016
ATPase/RNA helicase - hepatitis C virus (fragment)
C;Species: hepatitis C virus
C;Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 09-Jul-2004
C;Accession: S68016
R;Jin, L.; Peterson, D.L.
Arch. Biochem. Biophys. 323, 47-53, 1995
A;Title: Expression, isolation, and characterization of the hepatitis C virus ATPase/RNA
A;Reference number: S68016; MUID:96019946; PMID:7487072
A;Accession: S68016
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-386 <JIN>
C;Cross-references: UNIPROT:Q04045
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; nonstructural protein; nucleotide binding; P-loop; polyprotein
F;24-31/Region: nucleotide-binding motif A (P-loop)
F;86-91/Region: nucleotide-binding motif B
F;90-93/Region: DEXH motif

Query Match   5.7%; Score 871; DB 2; Length 386;
Best Local Similarity 46.3%; Pred. No. 1.5e-46;
Matches 187; Conservative 55; Mismatches 132; Indels 30; Gaps 8;

Qy   1119 PQYTAHATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASM 1178
                 :|:: .:||| |||| || |||||||||||| :||||:|: |||||||||||||
Db   1 PVFTDNSS--SPPVVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGF 58

Qy   1179 PKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYLT-GACSRN-YDVIICDECHATDAT 1236
                                                |||:|||||| |||||||||::|:|||||||
Db   59 YGKFLADGGCSGGAYDIIICDECHSTDAT 98

Qy   1237 TVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEEN 1296
                :||| ||||||||:|| |||||:|||||||||| :||:||||||||||||:|||||||||
Db   99 SILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTGEIPFYGKAIPLEA 158

Qy   1297 LKKGRHLIFEATKKHCDELANELARKGITAVSYYRGCDISKIP-EGDCVVVATDALCTGY 1355
                :|||||||:|||||||||||||:||:|| |||||||||||||| :||||||||||:||||
Db   159 IKGGRHLIFCHSKKKCDELATKLVALGINAVAYYRGLDVSVIPSSGDVVVVATDALMTGF 218

Qy   1356 TGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDG 1415
                :||||||||| :| |||| ||| |:||||||||| :||||||||||||||||| :::|||
Db   219 SGDFDSVIDCNTCVCTQTVDFSLDPTFTIETTLPQDAVSRTQRRGRTGRGKPGIYRFVAP 278

Qy   1416 SCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFS- 1474
                       .:| |::| || ::||:::|| :|::||||:||||||||||||
Db   279 GERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQDHLEFWEGVFTG 338

Qy   1475 MVNPEPSFVNTAKRTADNYVLLTAAQLCHQYGYAAPNDAPRW 1518
                   :.   .:| |: ||||::| || | :|:| :|| .|
Db   339 LTHIDAHFLSQTKQSGENFPYLVAYQATVCAR-AHAPP---PSW 378

Search completed: October 27, 2005, 15:49:25
Job time : 115 secs

Db   390 LVLYSPNNSGTTWLNRLLTTLPRSSCIPDSYF-QQADYCDKVSAMLRRLSLTRTVVALV 448
               .::: | ||||| :||||||||| :: ||||| :| :|:
Qy   1853 RWLHTPEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGP 1912

Db   449 N--REPKVDEV-QVGVVWDLWEWIMRQVRVVISRLRA----LCPVVSLPLWHCGEGWSGE 501
               .   :|:. : :.:.   : :: : |:| ||||:|||:|||| |:|
Qy   1913 WIGSGMLQARCPCGAELIFSVENGFAK--LYKGPRTCSNYWRGAVPVNARLCGSARPDPT 1970

Db   502 WLLDGHVESRCLCGVITGDVLNGQLKDVY-STKLCRHYWMGTVPVNMLGYGETSPLLA 560
               ||||| :|||| || || |:|:|:|: | ::|| ::: :: : |: |: :::|| :
Qy   1971 DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVA----VDGVQVQC 2026

Db   561 SDTPKVVPFGTSGWAEVVTPTHVVIRRTSCYKLLRQQI--LSAAVAEPYYVDGIPVS- 616
               ||||:: ||||: ||:| | :: ::| :| ::   :
Qy   2027 YLGEPKTWTTSACCYGP------DGKGKTVKLPFRVDGHTPG-VRMQLNLRDALETNDC 2079

Db   617 WEADARAP----AMVYGPGQSVTIDGERYTLPHQLRMRNVAPSEVSSEVSIEIGTEDS 672
                  ||.|.   :||::: :| | ::| |:::|: | :: :.|
Qy   2080 NSTN-NTPSDEAAVSALVFKQELRRTNQLLBAISAGVDTTKJ----------------P 2121

Db   673 ELTEADLPPAAAALQAI--ENAARILEPHIDVIMEDCSTPSLCGSSREMPVWEEDIPRTP 730
               |:| ::| ::::|:: :   ::| .|| : | :|::|| |:.|:|| |:|:::|.
Qy   2122 APSIEEVVRKRQFRARTGSLTLPPPRSVPGVSCPESLQRSDPLEGPSNL----------- 2172

Db   731 SPALISVTESSSDEKTPSASSSQEDTPSS-----DSFEVIQBSDTABSEESVFNVALSVLK 786
               |.|:| |::|||: |.|:.|::|:|||| |||: ::.||:: :::||||:|: ||:|:
Qy   2173 ---PPSPPVVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKKEVSEWSDES 2229

Db   787 ALFPQSDATRKLTVRMSCCVEKSVTRFFSLGLTVA------------DVASLCEME 830
               :|||| |:.|::   ::: ||:.| ||||| ||| ||:|:|:|
Qy   2230 WSTATTASSYVTGPPYPKIR---GKDSTQSAPAKRPTKKKLGKSEFSCSMSYTWTDVISF 2286

Db   831 IQNHTAYCDKVRTPLELQVGCLVGNELTFEC------DKCEARQETLASFSYIWSGVPLT 884
                 :|:||| :|| : :.||: |.|: | |::||:||::|:.|| |||
Qy   2287 KTASKVLSATRAITSGFLKQRSLVVYTEPRDAELRKQKVTINRQLFPPSY-HKQVRLAK 2345

Db   885 RATPAKPPVVRPVGSLLVADTTKVYVTNPDNVGRRVDKVTFWRAPRVHDKFLVDSIERAK 944
               ::||:||||| ||| |||||||||:: |||||||| ||:||:|||||||:| ||:|||
Qy   2346 EKASKVVGVVMDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLD-LQKCVERAGEIP 2404

Db   945 KSAQACLSMGYTYEEAIRTVRPHAAMGWGSKVSVKDLATPAGKMAVHDRLQEILEGTPVP 1004
               |:: ||::|||||::| ::.|:|::||:|| ::.| : :.| |:|:.:|::||: |::|
Qy   2405 SHYRQTVIVPKKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAVMGDA 2464

Db   1005 --FTLTV---KKEVFFK--DRKEEKAPRLIVFPPLDFRIAEKLILGDPG-RVAKAVLGGA 1056
                 ||:: |::.|| ::|::: :.||: ||.::::|:|::| . :| |||: |.
Qy   2465 YGF-VDPRTRYKRLLSMW-SPDAVGATC-DTVCFDSTITPEDIMVETDIYSAAKLSDQHR 2521

Db   1057 YAFQYTPNQRVREMLKLWESKKTPCAICVDATCFDSSITEEDVALSTELYA---LASDHP 1113
               ||||| |::| ::|::: |.|:: ||:|:|: |||::||||:| |.: | |||||
Qy   2522 AGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSNSLTCWLKVNAAAEQAGMKN 2581

Db   1114 EWVRALGK--YYASGTMVTPEGVPVGERYCRSSGVLTTSASNCLTCYIKVKAACDRVGLKN 1172
               |:||:||| |:||| ||| ||||| |:|:|| ||| .||.|||::|:|||: ||:||
Qy   2582 PRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVNGAPQDCVPQPKYSLEBLTSCSSNV 2641

Db   1173 VSLLIAGDDCLJICERPMCCDPSEALGRALASY------GYACEPSYHASLDTAPFCSTWL 1226
               .|::| ||||:| :|::| |||:|| |:|.: |:|: |:   ::||| ||| |
Qy   2642 TSGITKSGKPYYFLTRDPRIPLGRCSAEGLGY-NPSAAWIGYLIHYPCLWVSR-VLAVH 2699

Db   1227 AE-CNADGKRHFFLTTDFRRPLARMSSE---YSDPMASAIGYIL-LYPWHPITRWVIIPH 1281
               | ||: |:|||:| ||| ||| ||| |||.|||||||| | |||:|||| ::|
Qy   2700 FMEQMLFEDKLP-ETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEIlRVSQ 2758

Db   1282 VLTCAFRGGGTPSDPVWCQVHGNYYKFPLDKLPNIIVALHGPAALRVTADTTKTKMBEAGK 1341
               ||||| :||||||| ||| :.|||| |||||||| |: ||| |:| ||| ||| |:.|||
Qy   2759 SLTDMTMPPLRAWRKKARAVLASAKRGGAHAKLARFLLWHATSR-PLPDL-PPPEI 2808

Db   1342 VLSDLKLPGLAVHRKKAGA-LRTRMLRSRGWAELARCLLIWHPGLRLPPPEI 1391

RESULT 15

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.

- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

  http://www.pir.uniprot.org/database/archive.shtml

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

THIS PAGE BLANK (USPTO)

```
                    GenCore version 5.1.6
              Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:        October 27, 2005, 15:29:48 ; Search time 316 Seconds
                                  (without alignments)
                          4641.123 Million cell updates/sec

Title:    US-10-009-002-2
Perfect score:  15376
Sequence:  1 MPVISTQTSPVPAPRTRKNK.........KYLAVIVFALGLIAVGLiAiS 2864

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:    1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters:     1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   UniProt 03:*
             1: uniProt_sprot:*
             2: uniprot_trembl:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                                  SUMMARIES

Result          Query
No.     Score   Match  Length  DB  ID         Description
----    -----   -----  ------  --  --         -----------
1       15376   100.0   2864    2  Q9QEW5     Q9qew5 hepatitis g
2       15366    99.9   2864    2  Q999T0     Q999t0 hepatitis g
3       15317    99.6   2864    2  Q69422     Q69422 hepatitis g
4       10238    66.6   1925    2  Q8JKE4     Q8jke4 hepatitis g
5       3310.5   21.5   3033    2  Q7T7I0     Q7t7i0 hepatitis c
6       3306.5   21.5   3033    2  Q7T7I6     Q7t7i6 hepatitis c
7       3303.5   21.5   3033    2  Q7T7H9     Q7t7h9 hepatitis c
8       3303.5   21.5   3033    2  Q7T7I4     Q7t7i4 hepatitis c
9       3303.5   21.5   3033    2  Q9IZA1     Q9iza1 hepatitis c
10      3300.5   21.5   3033    2  Q7T7J1     Q7t7j1 hepatitis c
11      3296.5   21.4   3033    2  Q9QAX1     Q9qax1 hepatitis c
12      3293.5   21.4   3033    2  Q7T7J2     Q7t7j2 hepatitis c
13      3292.5   21.4   3019    2  Q68801     Q68801 hepatitis c
14      3289.5   21.4   3033    2  Q7T7H7     Q7t7h7 hepatitis c
15      3289.5   21.4   3033    2  Q7T7I7     Q7t7i7 hepatitis c
16      3288     21.4   3033    2  Q9IZA2     Q9iza2 hepatitis c
17      3287.5   21.4   3033    2  Q7T7I1     Q7t7i1 hepatitis c
18      3286.5   21.4   3033    2  Q7T7H6     Q7t7h6 hepatitis c
19      3285.5   21.4   3033    2  Q9DHD6     Q9dhd6 hepatitis c
20      3284.5   21.4   3033    2  Q7T7H8     Q7t7h8 hepatitis c
21      3283.5   21.4   3033    2  Q7T7I8     Q7t7i8 hepatitis c
22      3282.5   21.4   3033    2  Q7T7I2     Q7t7i2 hepatitis c
23      3282     21.3   3033    2  Q99IB6     Q99ib6 hepatitis c
24      3281.5   21.3   3033    2  Q7T7I3     Q7t7i3 hepatitis c
25      3281     21.3   3021    2  Q68870     Q68870 hepatitis c
26      3277.5   21.3   3033    2  Q9QF35     Q9qf35 hepatitis c
27      3276     21.3   3021    2  Q81258     Q81258 hepatitis c
28      3275     21.3   3008    2  Q9J3F4     Q9j3f4 hepatitis c
29      3273.5   21.3   3033    1  POLG_HCVJ8 P26661 h genome po
30      3273     21.3   3010    2  Q9QP06     Q9qp06 hepatitis c
31      3271.5   21.3   3033    2  Q7T7H3     Q7t7h3 hepatitis c
```

```
32      3265.5   21.2   3033    2  Q99IB8     Q99ib8 hepatitis c
33      3264.5   21.2   3032    2  Q99IB2     Q99ib2 hepatitis c
34      3264.5   21.2   3033    2  Q7T7H5     Q7t7h5 hepatitis c
35      3263.5   21.2   3033    2  Q7T7H4     Q7t7h4 hepatitis c
36      3263.5   21.2   3033    2  Q9IZA6     Q9iza6 hepatitis c
37      3263     21.2   3010    2  Q9J3H9     Q9j3h9 hepatitis c
38      3262     21.2   3010    2  Q9WMX2     Q9wmx2 hepatitis c
39      3262     21.2   3033    1  POLG_HCVJ6 P26660 h genome po
40      3261     21.2   3012    2  Q9WIK7     Q9wik7 hepatitis c
41      3260     21.2   3010    2  Q68826     Q68826 hepatitis c
42      3259     21.2   3010    2  Q9DTD6     Q9dtd6 hepatitis c
43      3258     21.2   3010    2  Q9DTE4     Q9dte4 hepatitis c
44      3257     21.2   3010    2  Q02829     Q02829 hepatitis c
45      3256.5   21.2   3010    2  Q9QP61     Q9qp61 hepatitis c
```

```
                                  ALIGNMENTS

RESULT 1
Q9QEW5
ID   Q9QEW5       PRELIMINARY;       PRT;   2864 AA.
AC   Q9QEW5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis GB virus B.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae.
OX   NCBI_TaxID=39113;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99434247; PubMed=10502525; DOI=10.1006/viro.1999.9941;
RA   Bukh J., Apgar C.L., Yanagi M.;
RT   "Toward a surrogate model for hepatitis C virus: An infectious
RT   molecular clone of the GB virus-B hepatitis agent.";
RL   Virology 262:470-478(1999).
DR   EMBL; AF179612; AAF01368.1; -.
DR   HSSP; P26664; 1HEI.
DR   MEROPS; S29.002; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; DEAD.
DR   InterPro; IPR001410; DEAD/DEAH_N.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR004109; PepEidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
```

```
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW   Transmembrane.
SQ   SEQUENCE   2864 AA;  312859 MW;  3CD167B2F1FE237 CRC64;

Query Match          100.0%;
Best Local Similarity 100.0%;   Score 15376;   DB 2;   Length 2864;   Pred. No. 0;
Matches 2864;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;


Qy    1 MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD   60

Qy   61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPV  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPV  120

Qy  121 CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY  180

Qy  181 CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG  240

Qy  241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK  300

Qy  301 LASQVPYAIATMPSSVHYLAVAGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LASQVPYAIATMPSSVHYLAVAGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI 360

Qy  361 AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF  420

Qy  421 KNNTWGCCRIRNVPSYCYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 KNNTWGCCRIRNVPSYCYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG 480

Qy  481 SKEMFKPHNWMSGHLYPEGSDTPIVVFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 SKEMFKPHNWMSGHLYPEGSDTPIVVFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ  540

Qy  541 GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY  600

Qy  601 LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL  660

Qy  661 LGFVPMAAGLPLTFFVAAAAAQPDYDWMWVRLLVAGLVLWAGRNRGHRIALLVGPWPLVAL 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 LGFVPMAAGLPLTFFVAAAAAQPDYDWMWVRLLVAGLVLWAGRNRGHRIALLVGPWPLVAL 720

Qy  721 LTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLMQRWENWF  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 LTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLMQRWENWF  780

Qy  781 WNVTLRPERFFLVLVCFPGATYDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 WNVTLRPERFFLVLVCFPGATYDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR  840

Qy  841 LGKCHAWYSHYVLKFPLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPPEGKARVY  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 LGKCHAWYSHYVLKFPLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPPEGKARVY  900

Qy  901 RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPFTLQCLSERGTLSAMAVV  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPFTLQCLSERGTLSAMAVV  960

Qy  961 MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020

Qy 1021 DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG 1080

Qy 1081 APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ 1140

Qy 1141 ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200

Qy 1201 TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT 1260

Qy 1261 ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA 1320

Qy 1321 RKGITAVSYYRGCDISKIPEGDCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT  1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 RKGITAVSYYRGCDISKIPEGDCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT  1380

Qy 1381 FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCCTPSGMVPECNIVEAFDAAKAWYG 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCCTPSGMVPECNIVEAFDAAKAWYG 1440

Qy 1441 LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNVVLLTAAQ 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNVVLLTAAQ 1500

Qy 1501 LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT 1560

Qy 1561 SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEECASFIP  1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEECASFIP  1620

Qy 1621 LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA 1680

Qy 1681 SCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAGTALGTWTSVGF  1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 SCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAGTALGTWTSVGF  1740

Qy 1741 VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT 1800

Qy 1801 AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGLLEASTPWSVISACIRWLHTPTE 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGLLEASTPWSVISACIRWLHTPTE 1860

Qy 1861 DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ 1920

Qy 1921 ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSSARPDPTDWTSLVVNYG 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921 ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSSARPDPTDWTSLVVNYG 1980

Qy 1981 VRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDGVQVQCYLGEPKTPWTTSAC 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  1981  VRDYCKYERMGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDGVQVCYLGEPKTPWTTSAC  2040
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE  2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE  2100

Qy  2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPRSVPGVSCPESL  2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPRSVPGVSCPESL  2160

Qy  2161  QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK  2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  QRSDPLEGPSNLPFSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK  2220

Qy  2221  EVSEWNSDESWSTATTASSYVTGPPYPKIRGDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  EVSEWNSDESWSTATTASSYVTGPPYPKIRGDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280

Qy  2281  TDVISFKTASKVLSATRAITSGFLKQRSLVYTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TDVISFKTASKVLSATRAITSGFLKQRSLVYTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340

Qy  2341  VRLAKEKASKVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  VRLAKEKASKVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400

Qy  2401  GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV  2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV  2460

Qy  2461  MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520

Qy  2521  RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEBQAGMK  2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEBQAGMK  2580

Qy  2581  NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN  2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN  2640

Qy  2641  VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHYPCLWVSRVLAVHF  2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHYPCLWVSRVLAVHF  2700

Qy  2701  MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL  2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL  2760

Qy  2761  TDMTMPPLRAWRKKARAVLASAKRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN  2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  TDMTMPPLRAWRKKARAVLASAKRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN  2820

Qy  2821  YCDVVSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS  2864
          ||||||||||||||||||||||||||||||||||||||||||||
Db  2821  YCDVVSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS  2864
```

RA   Sbardellati A., Scarselli E., Verschoor E., De Tomassi A., Lazzaro D.,
RA   Traboni C.;
RT   "Generation of infectious and transmissible virions from a GB virus B
RT   full-length consensus clone in tamarins.";
RL   J. Gen. Virol. 82:2437-2448(2001).
DR   EMBL; AJ277947; CAC33083.1; -.
DR   HSSP; P26664; 1HEI.
DR   MEROPS; U39.001; -.

```
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_NS5a1.
DR   InterPro; IPR002170; HCV_RdRP.
DR   InterPro; IPR004109; PepEidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_Dol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW   Transmembrane.
SQ   SEQUENCE  2864 AA;  312814 MW;  79EE4E6F9DBAA749 CRC64;
```

```
Query Match             99.9%;  Score 15366;  DB 2;  Length 2864;
Best Local Similarity   99.9%;  Pred. No. 0;
Matches 2860;  Conservative   4;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1  MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD   60

Qy     61  GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLVGPLVAGAVVRPV  120
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLVGPLVAGAVVRPV  120

Qy    121  CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY  180

Qy    181  CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVWGALVTCDALDIG  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVWGALVTCDALDIG  240

Qy    241  ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK  300
```

```
RESULT 2
Q99ST0
ID   Q99ST0      PRELIMINARY;      PRT;   2864 AA.
AC   Q99ST0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis GB virus B.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae.
OX   NCBI_TaxID=39113;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21446677; PubMed=11562537;
```

```
Qy   301  LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI   360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI   360

Qy   361  AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF   420
          ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db   361  AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPVTLEYNNSISWYPYTIPGARGCMVKF   420

Qy   421  KNNTWGGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG   480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  KNNTWGGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG   480

Qy   481  SKEMFKPHNWMSGHLYPEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ   540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  SKEMFKPHNWMSGHLYPEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ   540

Qy   541  GFYSDVKDLATGLIITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY   600
          |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db   541  GFYSDVKDLATGLIITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY   600

Qy   601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYLRCRLRYAAL   660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYLRCRLRYAAL   660

Qy   661  LGFVPMAAGLPLTFFVAAAAAQPDYDWWVRLLVAGLVLWAGNRGHRIALLVGPWPLVAL   720
          ||||||||||||||||||||||||||||||||||||||:|||||::|||||||||||||
Db   661  LGFVPMAAGLPLTFFVAAAAAQPDYDWWVRLLVAGLVLWAGRDRGHRIALLVGPWPLVAL   720

Qy   721  LTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQRWENWF   780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  LTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQRWENWF   780

Qy   781  WNVTLRPERRFFLVLVCFPGATYDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  WNVTLRPERRFFLVLVCFPGATYDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR   840

Qy   841  LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPEGKARVY   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPEGKARVY   900

Qy   901  RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPPTLQCLSERGTLSAMAVV   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPPTLQCLSERGTLSAMAVV   960

Qy   961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN  1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN  1020

Qy  1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG  1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG  1080

Qy  1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ  1140

Qy  1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN  1200

Qy  1201  TGASLITYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  TGASLITYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT  1260

Qy  1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA  1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA  1320

Qy  1321  RKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTCGPDSVYDCSLMVEGTCHVDLDPT  1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  RKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTCGPDSVYDCSLMVEGTCHVDLDPT  1380

Qy  1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYVVDGSCTPSGMVPECNIVEAFDAAKAWYG  1440
```

```
Db  1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYVVDGSCTPSGMVPECNIVEAFDAAKAWYG  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1441  LSSTEAQTILDYTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  LSSTEAQTILDYTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ  1500

Qy  1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT  1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT  1560

Qy  1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEBCASFIP  1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEBCASFIP  1620

Qy  1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA  1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA  1680

Qy  1681  SCVFAFIAGITTLPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAGTALGTWTSVGF  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  SCVFAFIAGITTLPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAGTALGTWTSVGF  1740

Qy  1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT  1800

Qy  1801  AGPDHWPNRLLIMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  AGPDHWPNRLLIMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE  1860

Qy  1861  DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ  1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ  1920

Qy  1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVNYG  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVNYG  1980

Qy  1981  VRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDGVQVQCYLGEPKTPWTTSAC  2040
          ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  VRDYCKYEKLGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDGVQVQCYLGEPKTPWTTSAC  2040

Qy  2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE  2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE  2100

Qy  2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL  2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL  2160

Qy  2161  QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK  2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK  2220

Qy  2221  EVSEWSDESWSTATTASSYVTGPPYPYKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  EVSEWSDESWSTATTASSYVTGPPYPYKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280

Qy  2281  TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340

Qy  2341  VRLLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  VRLLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400

Qy  2401  GEIPSHYRQTVIVPKEEVFVKTPQKPPRRLISYPHLEMRCVEKMYYGQVAPDVVKAV  2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  GEIPSHYRQTVIVPKEEVFVKTPQKPPRRLISYPHLEMRCVEKMYYGQVAPDVVKAV  2460

Qy  2461  MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520
```

```
Db   2461 MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH 2520
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2521 RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK 2580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK 2580

Qy   2581 NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN 2640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN 2640

Qy   2641 VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF 2700
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF 2700

Qy   2701 MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL 2760
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL 2760

Qy   2761 TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN 2820
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN 2820

Qy   2821 YCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS 2864
               |||||||||||||||||||||||||||||||||||||||||||
Db   2821 YCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS 2864
```

```
RESULT 3
Q69422
ID   Q69422;         PRELIMINARY;         PRT;         2864 AA.
AC   Q69422;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polypeptide.
OS   Hepatitis GB virus B.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae.
OX   NCBI_TaxID=39113;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95241511; PubMed=7724574;
RA   Simons J.N., Pilot-Matias T.J., Leary T.P., Dawson G.J., Desai S.M.,
RA   Schlauder G.G., Muerhoff A.S., Erker J.C., Buijk S.L., Chalmers M.L.,
RA   van Sant C.L., Mushahwar I.K.;
RT   "Identification of two flavivirus-like genomes in the GB hepatitis
RT   agent.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:3401-3405(1995).
DR   EMBL; U22304; AAC54059.1; -.
DR   HSSP; P26664; 1HEI.
DR   MEROPS; S29.002; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR004109; PepTidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW   Transmembrane.
SQ   SEQUENCE   2864 AA;   312702 MW;   5F5A7D8FAF0CDE81 CRC64;
```

```
Query Match             99.6%;
Best Local Similarity   99.6%;
Matches 2852;   Conservative    4;   Mismatches    8;   Indels    0;   Gaps    0;

Qy    1 MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD 60

Qy   61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPV 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPV 120

Qy  121 CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180

Qy  181 CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240

Qy  241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK 300

Qy  301 LASQVPYAIATMFSSVHYLAVGALIIYASRGKWYQLLALMLYIEATSGNPIRVPTGCSI 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LASQVPYAIATMFSSVHYLAVGALIIYASRGKWYQLLALMLYIEATSGNPIRVPTGCSI 360

Qy  361 AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPVTLEYNNSISWYPYTIPGARGCMVKF 420

Qy  421 KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKAILQYPG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKAILQYPG 480

Qy  481 SKEMFKPHNWMSGHLYFEGSDTPIVFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 SKEMFKPHNWMSGHLYFEGSDTPIVFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ 540

Qy  541 GFYSDVKDLATGLITKDXAWKNYQVLYSATGALSLTGVTTKAVVLILGLCGSKYLILAY 600
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  541 GFYSDVKDLATGLITKDXAWKNYQVLYSATGALSLTGVTTKAVVLILGLCGSKYLILAY 600

Qy  601 LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFCCYLRCRLRYAAL 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFCCYLRCRLRYAAL 660

Qy  661 LGFVPMAAGLPLTFVAAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALLVGPWPLVAL 720
          |||||||||||||||||||||||||||:|| ||||||||||:|:| |||||||||||||
Db  661 LGFVPMAAGLPLTFFVAAAAAQPDYDWWVRLLVAGLVLWAGRDRGPRIALLVGPWPLVAL 720

Qy  721 LTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQRWENWF 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 LTLLHLATPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQRWENWF 780
```

```
Qy   781 WNVTLRPERFFLVLVCFPGATYDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR  840
         ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db   781 WNVTLRPERFFLVLVCFPGATYDTLVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR  840

Qy   841 LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPFEGKARVY  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPFEGKARVY  900

Qy   901 RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPPTLQCLSERGTLSAMAVV  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPPTLQCLSERGTLSAMAVV  960

Qy   961 MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020

Qy  1021 DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDDYWCVCGALPMAVAKGSSG 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDDYWCVCGALPMAVAKGSSG 1080

Qy  1081 APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ 1140

Qy  1141 ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200

Qy  1201 TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT 1260

Qy  1261 ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA 1320

Qy  1321 RKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 RKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT 1380

Qy  1381 FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEARPDAAKAWYG 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEARPDAAKAWYG 1440

Qy  1441 LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ 1500

Qy  1501 LQLCHQYGYAAPNDAPRMQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 LQLCHQYGYAAPNDAPRMQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT 1560

Qy  1561 SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSTISVPTGATVAPVDEEIVEECASFIP 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSTISVPTGATVAPVDEEIVEECASFIP 1620

Qy  1621 LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA 1680

Qy  1681 SCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 SCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF 1740

Qy  1741 VFDMLGGYAAASSTACLLTFPKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 VFDMLGGYAAASSTACLLTFPKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT 1800

Qy  1801 AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE 1860
```

```
Qy  1861 DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ 1920

Qy  1921 ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDTDWTSLVVNYG 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDTDWTSLVVNYG 1980

Qy  1981 VRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVADGVQVQCYLGEPKTPWTTSAC 2040
         |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 VRDYCKYEKLGDHIFVTAVSSPNVCFTQVPPTLRAAVADGVQVQCYLGEPKTPWTTSAC 2040

Qy  2041 CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE 2100
         ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db  2041 CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSINNTPSDEAAVSALVFKQE 2100

Qy  2101 LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLLTLPPPPRSVPGVSCPESL 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLLTLPPPPRSVPGVSCPESL 2160

Qy  2161 QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGPDDLPSYPPKK 2220
         |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db  2161 QRSDPLEGPSNLPSSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGPDDLPSYPPKK 2220

Qy  2221 EVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLLGKSEFSCSMSYTW 2280
         |||||||:|||||:||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 EVSEWSDGSWSTTTTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLLGKSEFSCSMSYTW 2280

Qy  2281 TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ 2340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ 2340

Qy  2341 VRLAKEKASKVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 VRLAKEKASKVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA 2400

Qy  2401 GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV 2460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV 2460

Qy  2461 MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH 2520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH 2520

Qy  2521 RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521 RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK 2580

Qy  2581 NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN 2640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581 NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN 2640

Qy  2641 VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF 2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641 VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF 2700

Qy  2701 MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNABILRVSQSL 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNABILRVSQSL 2760

Qy  2761 TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN 2820

Qy  2821 YCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLLIAVGLAIS 2864
         ||||||||||||:|||||||||||||||||||||||||||||||
Db  2821 YCDVYSPEGDVFTTPQRRLQKFLVKYLAVIVFALGLLIAVGLAIS 2864
```

RESULT 4

Q8JKE4

ID   Q8JKE4   PRELIMINARY;   PRT; 1925 AA.

AC   Q8JKE4;

```
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Non-structural polyprotein.
OS   Hepatitis GB virus B.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae.
OX   NCBI_TaxID=39113;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22092527; PubMed=12097587;
RX   DOI=10.1128/JVI.76.15.7736-7746.2002;
RA   De Tomassi A., Pizzuti M., Graziani R., Sbardellati A., Altamura S.,
RA   Paonessa G., Traboni C.;
RT   "Cell clones selected from the Huh7 human hepatoma cell line support
RT   efficient replication of a subgenomic GB virus B replicon.";
RL   J. Virol. 76:7736-7746(2002).
DR   EMBL; AJ428955; CAD21957.1; -.
DR   HSSP; P26664; 1HEI.
DR   MEROPS; U39.001; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Polyprotein.
SQ   SEQUENCE   1925 AA;  208751 MW;  5B7D62841311 9C72 CRC64;

Query Match     66.6%;   Score 10238;   DB 2;   Length 1925;
Best Local Similarity 99.9%;   Pred. No. 0;
Matches 1922   Conservative    2;   Mismatches    2;   Indels    0;   Gaps    0;
```

```
Qy  941 APFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSK 1000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2 APFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSK 61

Qy 1001 GRRLAHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSD 1060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 GRRLAHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSD 121

Qy 1061 DPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQ 1120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122 DPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQ 181

Qy 1121 YTAHATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPK 1180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  182 YTAHATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPK 241

Qy 1181 YMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLG 1240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  242 YMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLG 301
```

```
Qy 1241 IGKVLTEAPSKNVRLVLVLATATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKG 1300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  302 IGKVLTEAPSKNVRLVLVLATATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKG 361

Qy 1301 RHLIFEATKKHCDELANELARKGITAVSYYRGCDISKIPEGDCVVATDALCTGYTGDFD 1360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  362 RHLIFEATKKHCDELANELARKGITAVSYYRGCDISKIPEGDCVVATDALCTGYTGDFD 421

Qy 1361 SVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYVVDGSCTPS 1420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  422 SVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYVVDGSCTPS 481

Qy 1421 GMVPECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEP 1480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  482 GMVPECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEP 541

Qy 1481 SFVNTAKRTADNYVLLTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADAC 1540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  542 SFVNTAKRTADNYVLLTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADAC 601

Qy 1541 PGPEPSEVTRYQMCFTEVNTSGTAALAVGVAMAYLAIDTFGATCVRRCWSITSVPTGA 1600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  602 PGPEPSEVTRYQMCFTEVNTSGTAALAVGVAMAYLAIDTFGATCVRRCWSITSVPTGA 661

Qy 1601 TVAPVVDEEIVEECASFIPLEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAAT 1660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  662 TVAPVVDEEIVEECASFIPLEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAAT 721

Qy 1661 ILAIIEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGAL 1720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  722 ILAIIEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGAL 781

Qy 1721 AFMMAGAAGTALGTWTSVGFVPDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAF 1780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  782 AFMMAGAAGTALGTWTSVGFVPDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAF 841

Qy 1781 NPAAGVVGVLSACAMPALTTAGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILE 1840
         |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db  842 NPAAGVVGVLSACAMFALTTAGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILE 901

Qy 1841 ASTPWSVISACIRWLHTPTEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGC 1900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  902 ASTPWSVISACIRWLHTPTEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGC 961

Qy 1901 PFYSCQKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNAR 1960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  962 PFYSCQKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNAR 1021

Qy 1961 LCGSARPDPTDWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAAVD 2020
         |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db 1022 LCGSARPDPTDWTSLVVNYGVRDYCKYEKLGDHIFVTAVSSPNVCFTQVPPTLRAAAVD 1081

Qy 2021 GVQVQCYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCN 2080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1082 GVQVQCYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCN 1141

Qy 2081 STNNTPSDEAAVSALVKQELRRTNQLLEAISAGVDTTKLPAPSIEEVVVRKRQFRARTG 2140
         ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db 1142 STNNTPSDEAAVSALVFKQELRRTNQLLEAISAGVDTTKLPAPSIEEVVVRKRQFRARTG 1201

Qy 2141 SLTLPPPPRSVPGVSCPESLQRSDPLEGPSNLPPSPPVLQAMPMPLLGAGECNPFTAIG 2200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1202 SLTLPPPPRSVPGVSCPESLQRSDPLEGPSNLPPSPPVLQAMPMPLLGAGECNPFTAIG 1261

Qy 2201 CAMTETGGGPDDLPSYPPKKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAK 2260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1262 CAMTETGGGPDDLPSYPPKKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAK 1321

Qy 2261 RPTKKKLGKSEFSCSMSYTWTDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAEL 2320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1322 RPTKKKLGKSEFSCSMSYTWTDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAEL 1381

Qy 2321 RKQKVTINRQPLFPPSYHKQVRLAKEKASKVVGVWMDYDEVAAHTPSKSAKSHITGLRGT 2380
```

```
Db  1382 RKQKVTINRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGT 1441
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2381 DVRSGAARKAVLDLQKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLE 2440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1442 DVRSGAARKAVLDLQKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLE 1501

Qy  2441 MRCVEKMYYGQVAPDVVKAVMGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTIT 2500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1502 MRCVEKMYYGQVAPDVVKAVMGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTIT 1561

Qy  2501 PEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSS 2560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1562 PEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSS 1621

Qy  2561 SNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAP 2620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1622 SNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAP 1681

Qy  2621 QDCVPQPKYSLEELTSCSSNVTSGITKSGKPYFLTRDPRIPLGRCSAEGLGYNPSAAWI 2680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1682 QDCVPQPKYSLEELTSCSSNVTSGITKSGKPYFLTRDPRIPLGRCSAEGLGYNPSAAWI 1741

Qy  2681 GYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGI 2740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1742 GYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGI 1801

Qy  2741 EAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHA 2800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1802 EAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHA 1861

Qy  2801 TSRPLPDLDKTSVARYTTFNYCDVYSPEGDVFTPQRRLQKFLVKYLAVIVFALGLIAVG 2860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1862 TSRPLPDLDKTSVARYTTFNYCDVYSPEGDVFVTPQRRLQKFLVKYLAVIVFALGLIAVG 1921

Qy  2861 LAIS 2864
         ||||
Db  1922 LAIS 1925


RESULT 5
Q7T7I0
ID   Q7T7I0;      PRELIMINARY;      PRT;      3033 AA.
AC   Q7T7I0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tanabe Y., Nagayama K., Enomoto N., Izumi N., Tazawa J., Sato C.,
RA   Watanabe M.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY232742; AAP55697.1; -.
DR   HSSP; Q8JYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH N.
DR   InterPro; IPR002521; HCV_capsid.
DR   InterPro; IPR002519; HCV_core.
DR   InterPro; IPR002531; HCV_env.
DR   InterPro; IPR000745; HCV_NS1.
DR   InterPro; IPR001490; HCV_NS4a.
DR   InterPro; IPR002868; HCV_NS4b.
DR   InterPro; IPR002166; HCV_NS5a.
DR   InterPro; IPR004109; HCV_RdRP.
DR   InterPro; IPR009003; Peptidase_S29.
DR   InterPro; IPR002518; Pept_Ser_Cys.
DR   InterPro; IPR007095; Pept_U39_HCV_NS2.
DR   InterPro; IPR007094; RNA_pol_DS_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE  3033 AA;  330368 MW;  46C9F5030F039B12 CRC64;

Query Match             21.5%;  Score 3310.5;  DB 2;  Length 3033;
Best Local Similarity   30.1%;  Pred. No. 5.8e-206;
Matches  986;  Conservative  449;  Mismatches 1171;  Indels  669;  Gaps 111;

Qy   8 TSPVPAPRTRKNKQTQASYPVSIK--------------------TSVERGQRAKRKVQRDARPR 51
         ||| :|| :|    :                            :: :
Db   3 TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR 59

Qy  52 NYKIAGIHDGLQTLAQAALPAH---------GWGRQDPRHKSRNL 87
         : ::                          |||||:|:||:|||
Db  60 GRRQPIPKDRRSTGKSWGKPGYPWPLYGNEGCGWAGWLLSPRGSRPTWGPTDPRHRSRNL 119

Qy  88 GILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLFV 144
         :| |   :: ||                               ::        ||||
Db 120 GRVIDTITCGFADLMGYIPVVGAPV-GGVARALAHGVRVLEDGINYATGNLPGCSFSIFL 178

Qy 145 VCLLS-LACPCSGARVTDPDTNTTILNCCQRNQVIYCSPSTCLHEPGCVICADE----- 198
         ::  |  :|:                                      :
Db 179 LALLSCVTVPVSAVEVRN-ISSSYYATNDCSNNSITWQLTNAVLHLPGCVPCENDNGTLR 237

Qy 199 CWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV---- 254
         |: |      : ::  :: ||  |:  |::  |::  : : |   |  :||:
Db 238 CWIQVTPNVAVKHRGALTHN-LRTHVDMIVMAATVCSALYVGDVCGAVMLVSQALIVSPE 296

Qy 255 RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGMAGKVEAVIFLTKLASQVPYAIATMFS 314
                                 |  |  :  : :    :  :  ||
Db 297 RH------NFTQECNCSIYQGQITGHRMAWDMMLNWSPTLTMILAYAARVPELVLEVVF 349

Qy 315 SVHYLAVGALIYYASRGKWYQLLLALMLY-IEAT---SGNPIRVPTGCSIAEFCSPLMI 369
         : | ::  :  :|  :: : :: ::|||    :  || ||:||||  |:|  |  |
Db 350 GGHWGVMFGLAYFSMQGAWAKVIAILLLVAGVDATTYSSGTAV---GSTLSSFAG-LFT 404

Qy 370 PCP------------------------CHSYLSENVSEVICYSPKWT-------------RPI- 395
        |                          ::|: ::|| :|
Db 405 RGPQQQLHLINTNGSWHINRTALNCNDSLQTGFIASLFYAKRFNSSGCPERLSSCRALDD 464

Qy 396 -----TLEYNNSIS---------W-YP------------------YTIPGARGCM----- 417
             ||:|:|: |                            :|  | ||:|:|
Db 465 FRIGWGALEYETNVTNDEDMRPYCWHYPPKPCGIVPARTVCGPVYCFTPSPVVVGTTDRQ 524

Qy 418 -------------VKFKNNT--------WGC----------------CRIR---NVPSYCT 438
                      | ||:|          |||                ||
```

```
Db    525 GVPTYSWGENETDVFLLNSTRPPRGAWFGCTWMNGTGFTKTCGAPPCRIRRDFNSTLDLL 584
            ||| |:  | ||| |::     ::|  ||  | | | :: |    | :  ::::
Qy    439 MGTDAVWNDTRNTYEACGVTPWLT-------TAWHNGSALKLAILQ---YPPGSKEMFK 486

Db    585 CPTDCFRKHPDTTYIKCGAGPWLITPRCLIEYPYRLMHYPCTVNFTIFKVRMYVGGVE--- 641
            |      |: |||      |   |:|::: | |    |
Qy    487 PHNWMSGHLYF---------EGSDTPIVYFYDPVNSTLLPPERWARLPGTPVVRGS 534

Db    642 --HRMSAACNFTRGDRCRLEDRDRGQQSPLLH---------ST----TEWAVLPCS--- 682
                       ||: |||    :|                         ||:  :|
Qy    535 WLQVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVLILLGLCGSK 594

Db    683 -------FSDLPALSTGLMHLHQNIVDVQYLYGLSPAVTRYIVKWEWVLLFLLLADAR 734
                 |::  ||    :|        | ||  ||  ||  | :|:|  :||| | |
Qy    595 YLILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPFALIFFICCYLRCR 654

Db    735 --VCACLWMLI----------ILGQAEAALEKLI--------- 757
             :|||  |:|             |:  | ||
Qy    655 LRYAALLGFVPMAAGLPLTFVAAAAQPDYDWWRLLVAGLVLWAGRNRGHRIALLVGP 714

Db    758 -----LHSASASANGPLWFPIFFTTA-----WYLKGRVVPVATYS---------VLGL 797
                 | || |:| |  | | ::::|     ||:|||:||||||
Qy    715 WPLVALLTLLHLVTPASAFDTEIIG--GLTIPPVVALVVMSRFGFFAHLLPRCALVNSYL 772

Db    798 WSF--LLLVLALPQQAYALDTAEQGELGLVILAIISIFTLT--PAYKILLSRSVVWLSYM 853
             |        | |    :|  :|         ||| |  :|  ||||||||||||||||
Qy    773 W-------ENWFW-NVTLRPERPFLV---LVCPPGATYDALVTFCVC 811

Db    854 LVLABAQIQQWVPPLEARGGRDGIIWAAVILHPRLVFEVTKWLLAILGSAY-------- 904
             | |      || :|| |   ||   |  :| |   || |||||| |||
Qy    812 HVALLCLTSSAASFGTDSRVRAHRM-----LVRLGKCHAWYSHYVLKFFLLVFGE-NGV 865

Db    905 --LLKASLLRVPYF---VRAHALLRVCTLVR-----HLAGAKYI-QMLLITLGRWTGT 951
              ||||| |||||    |||| || |||||     |||| |||||||||||||||||
Qy    866 FFYKHLHGDVLPNDFASKLPLQ----------EP--FFPPEGKARVYRNEGRRLAC 909

Db    952 ---------PLSTWAAQGLRDLAVAVEPVVFSPMEKKVIWGAE--TVAC 997
                     ||||||||:|||||| |||||||||||||||||||  ||||
Qy    910 GDTVDGLPVVARLGDLVFAGLAMPPD-----GWAITAPFTLQCLSERGTLSAMAVVMTGI 964

Db    998 GDILHLGLPVSARLGKEVLLG---PADSYTSKGWKLLAPITAYTQQTRGLLGAIVVSLTGR 1054
            ||  | |||||||||||||||   |  ||| |||  ||| ||:||||||||:|:|||:|||
Qy    965 DPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDI 1024

Db   1055 DKNEQAGQVQLSSVTQSFLGTSISGVLWTVYHGAGNKTLAGPKGPVTQMYTSAEGDLVG 1114
            |||| || ||||| ||||||| |||||||||||||||||||||||||||||| || ||
Qy   1025 YQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPIL 1084

Db   1115 WPSSPPGTKSLDPCTCGAVDLYLVTRNADVIPVRRKDRRGALLSPRPLSTLKGSSGGPVL 1174
             | |||||||||||||||||||||| |  |||:||||||||||||||||||||||||||
Qy   1085 CSSGHVIGMFTA---ARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQI 1141

Db   1175 CPRGHAVGLFRAAVCARGVAKSIDFVPVESLAIATRTPSFSDNS---APPAVPQTYQVGY 1231
             | ||:|| | || ||||||| ||||||||||||||:|||:||:   |||||||| ||||
Qy   1142 LIAPTGSGKSTKLPLSYMQBKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNT 1201

Db   1232 LHAPTGSGKSTKVPAAYASQGYKVLVLNPSVAATLGFGAYMSKAHGINPNIRTGVRTVT 1291
            | |||||||||| | ||| | ||||||||||| | || ||||| ||||| |||| | |
Qy   1202 GASLTYSTYGMYL-TGACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLA 1259

Db   1292 GDPITYSTYGKFIANGGCSAGAYDVIICDECHSVDATTILGIGTVLDQAETAGARLVVLA 1351
            || ||||||| | | ||||||||||||||||| |||||||||| |  | ||| ||||||
Qy   1260 TATPPGVIPTPHANTEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKHCDELANEL 1319

Db   1352 TATPDGVTVTTPHSNIEEVALGHEGEIPFYGKAIPLAFIRGGRHLIFCHSKKKKKCDELAAAL 1411
            ||||:||  | || |||: | | |||||:||| | |||||||||||| | |||| |||| ||
Qy   1320 ARKGITAVSYYRGCDISKIP-EGDCVVATDALCTGYTGDFDSVVDCSLMVEGTCHVDLD 1378

Db   1412 RGMGVNAVAYYRGLDVSVIPTQGDVVVVATDALMTGYTGDFDSVIDCNVAVTQIVDFSLD 1471
              :| ||| |||| |:|:||:||||| |||||| |||||||||||:|| || ||:||||||
Qy   1379 PTFTMGVRCVCGVSAIVKGQRRGTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAW 1438

Db   1472 PTFTITTQTVPQDAVSRSQRRGRTGRGRLGIYRYVSSGERPSGMFDSVVLCECYDAGAAW 1531
            |||| ||| || |||||||||||||||| |||:|| ||||||||||||||||||||||||
Qy   1439 YGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLT 1497
```

```
Qy   1439 YGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLT 1497
            ||  |  || || || | ||||| |||||||||||  ||||||||||||||||||||||
Db   1532 YELTPAETTVRLRAYFNTPGLPVCQDHLEFWEAVFTGLTHIDAHFLSQTKQGGDNFAYLT 1591

Qy   1498 AAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADAC---PGPEP- 1545
            | || |     |  | |  |  |   |:| | :|     |
Db   1592 AYQATVCAR---AKAPPSWD-VMWK-----CLTRLKPTLTGPTLLYR 1631

Qy   1546 -------SEVTRY-----QMCFTEVNTSGTAALAVGVAMAYLAIDFPGATCVRRC 1590
                   | |      :| | |  | :|  ||| || || | |  | |  | :|
Db   1632 LGAVTNEVTLTHPVTKYIATCMQADLEIMTS-TWVLAGGVLAAVAAYCLATGCVSIIGRI 1690

Qy   1591 WSITSVPTGATVAPVDEEEI-----VEECASFIPL-EAMVAAIDKLKSTI-------TT 1637
             :   ||| ||| | |:     :|::| |     | || | |||:| :          ||
Db   1691 HLNDQV---VVAP--DKEVLYEAFDEMEECSSKAALIEEGQRIAEMLKSKIQGLLQQAT 1744

Qy   1638 TSPFTLETALE-----KLNTFLGPHAATILAIIEYCCGLVTLPDNPFASCVFAPIAGITTP 1693
            |:      | |     :| |:| | |   | ||   | |:||::||| | | |  | | |:
Db   1745 RQAQDIQPAIQSSWPKLEQFWAKHMWNFISGIQYLAGLSTLPGNPAVASMMAFSAALTSP 1804

Qy   1694 LPHKIKMFLSLFGGAIASKLTDARGALAFMNAGAAGTALGTWTSVG---FVFDMLGGYAA 1750
            || | | |:| |||: || : |  ||  :: |  |:| ||   | |   : ||:| ||:|
Db   1805 LPTSTTILLNIMGGWLASQIAPPAGATGFVVSGLVGAAVG---SIGLGKILVDILAGYGA 1861

Qy   1751 ASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTTAGPD--HWPN 1808
              |:|   ::| :| |   ::: || |:| | | |:||:: ||  | :|  |||:  ||
Db   1862 GISGALVAFKIMSGEKPSVEDVVNLLPAILSPGALVVGVICAAILRRHVGGQGEGAVQWMN 1921

Qy   1809 RLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTEDDCGLI-- 1866
            || :::|| |:| |  : |::||  :::  :| |:| |:: ::|:  | |::  |
Db   1922 RLIAFASRGNHVAPTHYVAESDASQRVMQMLSSLT---ITSLLRRLHTWITEDCPVPCCS 1977

Qy   1867 -AWGLEIWQYVCNFPVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQARCPC 1925
               | | || |: | :| || : :     ::|| | ||||||| | || ||::  ||||
Db   1978 GSWLRDVWDWVCSILIDFKNWLSS-KLLPKMPGLPFISCQKGYRGVWAGTGVMTTRCPC 2035

Qy   1926 GAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARP-DPTDWTSLVVNYGVRDY 1984
            ||  : | ||    : |||| | | | | ||:| ||||    |     : :  | |
Db   2036 GANISGHVRMGTMKI-TGPKTCLNLWQGTFPINCYTEGPCVPKPPNYKTAIWRVAASEY 2094

Qy   1985 CKYEKMGDHIFVTAVSSPNVCFTQVP---PTLRAAVAVDGVQVQCYLGEPKTPWTTSACC 2041
             |       :|   || |:| | |      ||  |  ||||:    :|||  |   ||
Db   2095 VEVTQHGSFSVYTGLTSDNL---KVPCQVPAPEFFSWVDGVQIHRFAPIP------- 2141

Qy   2042 YGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSD---EAAVSALVFK 2098
            :||  |  ||| | || |:||      |||| | ||   |       |   ||| | ||
Db   2142 -GPFFRD---EVTFTVGLNSLLVVGSQLPCDPEPDTEVLASMLTDPSHITABAAARRLARG 2197

Qy   2099 QELRRTNQLLEAISA-----GVDTTKLPA- 2122
            : |::|    |    |     |  |   |
Db   2198 SPPSQASSSASQLSAPSLKATCTTHRMAYDCDMVDANLFMGGDVTRIESDSKVIVLDSLD 2257

Qy   2123 -------PSI-EEVVVRKRQFRARTGSLTLPP--PPRSVPGVSCP--ESLQRSDPLE 2167
                   |   ||    ||:    |  | |    ||  |||:       |:|::|
Db   2258 TMTEVEDDREPSVPSEYLVRKRKF---------PPALPPWARPDYNPPVIETWKR----- 2303

Qy   2168 GPSNLPPS---------PPVLQLAMPMPLLGAGECNPFTAIGCAMTE------- 2205
             |   |||         ||    ::|:|         ||  |
Db   2304 -PGYEPPTVLGCALPPTPQTPVPPPRRRRAKVLTQDNVGEVLREMADKVLSPLQDHNDSG 2362

Qy   2206 -------TGGGPDDLPSYPPKKEVSEWSDESWSTATTASSYVTGPPYPYKIRGK----- 2251
                   | || |    : |:|: | | ||| |  |::::       ||
Db   2363 HSTGADTGG-----TAASEAGSSSSMPPLEGEGPGDPDLEF 2406

Qy   2252 DSTQSAPAKR----------PTKKKLGKSEFSCSMSYTWTD-VISFKTASKVLSATR 2297
              ::| |       | | | | |  ||||||  || ||  | || | :   ||
Db   2407 EPAESAPPSEGECEVIDSDSKSWATVSDQEDSVICCSMSYSWTGALITPCGPEEEKLPIN 2466

Qy   2298 AITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQVRLAKEKASKVVGVWMD 2357
              |:  :::|   |:|   |||  | |||:   : |  |:|:|||  |::  |||  |
Db   2467 PLSNSLMRFHNKVYSTTSRASALRAKKVTFDRVQVLDAHYDSVLQDVDVKRAASVSARLLS 2526
```

```
Qy  2358  YDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEAGEIPSHYR-QTVIVPKE  2416
              :|||  ||||||        |||  :|:|         :||:  ||  ||  :::|
Db  2527  IEEACALTPPHSAKSRY-GFGAKEVRS-LSRRAVNHIRSVWEDLLEDQHTPIDTTVMAKN  2584

Qy  2417  EVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYGQVAPDVVKAVMGDAYGF-VDPRTRVK   2475
              |||  ||||| || ||||  || ||  |||||||     ||||    |    ||||:
Db  2585  EVFCVDPAKGGKKPARLLVYPDLGVRVCEKMALYDIAQKLPKAIMGPSYGFQYSPAERVD  2644

Qy  2476  RLLSMW--SPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYA  2533
              ||||    |    |||| |||||| |  ||| |  |:|| ||| ||  || |:|  :|
Db  2645  FLLKAWRSKKDPMGFSYDTRCFDSTVTERDIRTEESIYQACSLPEEARTVIHSLTERLYV  2704

Qy  2534  GGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTV  2593
              ||||    | |||||||:||| ||| ::|::||::|:||| | |:| | ||:|||| :|
Db  2705  GGPMTNSKGQSCGYRRCRASGVFTTSMGNTMTCYIKALAACKAAGIMDPMMLVCGDDLVV  2764

Qy  2594  IWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSNVTSGITKSGKPYY  2653
               :  :|  :|    ||  :|       | || |:| |::||:||||||   :   ||| 
Db  2765  ISESQGNEEDERNLRAFTEAMTRYSAPPGDLPRPEYDLELITSCSSNVSVALDSRGRRRY  2824

Qy  2654  FLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPET  2713
              ||||||  |   :||  : ::  ||   |:|  || | | |:      :   :   |:
Db  2825  FLTRDPTTPITRAAWETVRHSPVNSWLGNIIQYAPTIWVRMVIMTHFFSILLAQDTLNQN  2884

Qy  2714  VTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRK  2773
               ::: ||  |:| | |||:||   ||::||| : |::  |  ||  :  | |||||||:
Db  2885  LNFEMYGAVYSVNPLDLPAIIERLHGLDAFSLHTYSPHELSRVAATLRKLGAPPLRAWKS  2944

Qy  2774  KARAVLASAKRRGGAHAKLARFLL-WHATSR-----PLPDLDKTSVARYTTFNYCDVYSPE  2828
              :||| :||    || :|   |:|:|| ::|:     |||:  | : ||||    |||   ::
Db  2945  RARAVRASLIAQGGRAAICGRYLFNWAVKTKLKLTPLPEASRLDLSGWFT-----VGAGG  2999

Qy  2829  GDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAI  2863
              || :        :||  :|||:|  :  |:||:|
Db  3000  GDIYHSVSHARPRLL------LLCLLLSVGVGI   3027


RESULT 6

Q7T7I6
ID   Q7T7I6       PRELIMINARY;     PRT;    3033 AA.
AC   Q7T7I6;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tanabe Y., Nagayama K., Enomoto N., Izumi N., Tazawa J., Sato C.,
RA   Watanabe M.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY233736; AAP55691.1; -.
DR   HSSP; Q8JYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3033 AA;  329956 MW;  D58C8C087216598C CRC64;


Query Match             21.5%;  Score 3306.5;  DB 2;  Length 3033;
Best Local Similarity   30.3%;  Pred. No. 1.1e-205;
Matches  994;  Conservative  428;  Mismatches 1182;  Indels  673;  Gaps  110;

Qy    8  TSPVPAPRTRKNKQTQASYPVSIK---------------------------------TSVERGQRAKRKVQRDARPR   51
             |:|| ||   |  ||  | ||||:|                                 :|| || :|:|||:|
Db    3  TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR   59

Qy   52  NYKIAGIHDGLQTLAQAALPAH---------------------------------GWGRQDPRHKSRNL   87
                                                                       ||||:||||||:||
Db   60  GRRQPIPKDRRSTGKSWGKPGYPWPLYGNEGCGWAGWLLSPRGSRPTWGPTDPRHRSRNL  119

Qy   88  GILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG--WFGVHLFV  144
             | :| | :| :|   |||||||:|   |  ||| | ||:|||:| |||    |  :|:
Db  120  GKVIDTVTCGFADLMGYIPVVGAPV-GVARALAHGVRVLEDGINYATGNLPGCSFSIFL  178

Qy  145  VCLLS-LACPCSGARVTDPDTNTTILNCCQRNQVIYCSPSTCLHEPGCVICADE----  198
             : ||  | | |: | |  :|:  ::   :||:   |    ||  ||| :|
Db  179  LALLSCVTVPVSAVEVRN-ISSSYYATNDCSNNSITWQLSGAVLHLPGCVPCENDNGTLR  237

Qy  199  CWPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLVRHWL  258
             || :   : | :  ||: ::| ::::|:|| ||| || : ||  ||| ::| ::  |
Db  238  CWIQVTPNVAVKYHGALTHN-LRTHVDMIVMAATVCSALYVGVDVCGAVMIVSQAFIISPQ  296

Qy  259  IHIDLNETGTCYLEVPTGIDPGPLGFIGWMAGKVEAVIFLTKLASQVPYAIATMFSSVHY  318
             :| :|     || | ||   :  |     ||  :  :  :  | | |   | :: :  |
Db  297  HH---NFTQECNCSIYQGRITGHRMAWDMLNWSPTLTMILAYAARVPELALELVFGGHW   353

Qy  319  LAVGALIYASRGKWYQLLLALMLYI----------EATSG-------NP  351
             : :| ::|  :| :   :  |  :                         ||
Db  354  GVVFGLAYFSMQGAWAKVIAILLLVAGVDADTYASGAQAGRVTSGFASIFSPGAKQNINL  413

Qy  352  I------RVPTGCS-----------IAEF----CSPLMIPC------  371
             |      | | |             : :|     |  :  |
Db  414  IKTNGSWHINRTALNCNDSLNTGFIASLFYVQHFNSSGCPERMSSCRGLDDFRIGWGTLE  473

Qy  372  ----------------PCHSYLSENV-SEVICYSPKWTRPITLEYNN-----SIS  404
                              |   |  |:| :  :|  |    | :| |
Db  474  YETNVTNDEDMRPYCWHYPPKPCGIVPARTVCGPVYCFTPS---PVVVGTTDRQGAPTYS  530

Qy  405  WYPYTIPGARGCMVKFKDNT---------WGC-------CRIR---NVPSYCT  438
             |        || ||:| :|          |          |:|    ||  ||
Db  531  W------GENETDVFLLNSTRPPRGAWFGCTWMNGTGFTKTCGAPPCRIRRDYNSTLDLL  584

Qy  439  MGTDAVWNDTRNTYEACGVTPWLT------TAWHNGSALKLAILQ--YPGSKEMFK  486
```

```
Db   585  ||    ||||                                ||     ||||
          CPTDCPRKHPDATYIKCGAGPWLTPRCLVDPYPRLMHYPCTVNFTIFKVRMYIGGVE---  641
Qy   487  PHNWMSGHLYF----------EGSDTPIVFYFYDPVNSTLLPPERWARLPGTPPVVRGSW   535
             :    :      ||    ||||        ST----TEWAVLPCS------
Db   642  -HRFSAACNFTRGDRCGLEBDRDRGQQSPLLH-------ST----TEWAVLPCS------   682
Qy   536  LQVPQGFYSDVKDLATGLIITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKY  595
             :  :  ||:     ||
Db   683  -------FSDLPALSTGLHLHQNIVDVQVLYGLSPAVTRYIVKWEWVVLLFLLLADAR-   734
Qy   596  LILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYIRCRL   655
                             :|:|   ||
Db   735  -ICACLWMLI----------------ILGQAEAALEKLII----------L           758
Qy   656  RYAALLGFVPMAAGLPLTFVAAAAAQPDYDWWRLLVAGLVLWAGRNRGHRIALLVGPW     715
             :                         :    ||   ::::
Db   759  HSAS----AASANGL-LWFFIFFTAA------WYLKGRVVPVATYS------VLGLW      798
Qy   716  PLVALLTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLM--   773
              :|||||:           |:  |:|:|
Db   799  SF-LLIVLALPQQAYALNVAEQGELGLVVIVIISIFTLTPAYKILLSRSVVWLSYMLVL    856
Qy   774  -----QRW-----ENWFW-NVTLRPERFFLV--LVCFPGATYDALVTFCVCHVA         814
             : :
Db   857  AEAQIQQWMPVPPLEARGGRDGIIWVAVILPHRLVFEVTKWLLAILGSAY----------   904
Qy   815  LLCLTSSAASFFGTDSRVRAHRM------LVRLGKCHAWYSHYVLKFFLLVFGENGVFFYK  869
             :: :                 ||||              ::|:
Db   905  LLKASLLRVPYF------VRAHALLRVCTLVR-----HLAGARYIQMLLITVGRWTGTYIYD 955
Qy   870  HLHGDVLPNDPASKLPLQ----------EP--FFPFEGKARVYRNEGRRLACGDTV       913
             |||                     ||:  |
Db   956  HLS----------PLSTWAAQGLRDLAVAVEPVVFSPMEKKVIVWGAE--TVACGDIL    1001
Qy   914  DGLPVVARLGDLVFAGLAMPPD------GWAITAPFTLQCLSERGTLSAMAVVMTGIDPRT   968
             ||| |:                :        |:||:      ::| ||
Db  1002  HGLPVSARLGREVLLG---PADSYTSKGWKLLAPITAYAQQTRGLLSAIVVSLTGRDKNE  1058
Qy   969  WTGTIFRLGSLLATSYMGFVCDNVLVTYAHHGSKGRRLAHPTGSIHPITVDAANDQDIYQPP 1028
             ||:  :|||  :: : | |  :|: :|||||: | ||||:| ||
Db  1059  QAGVQVQVLSSVTQSFLGTSISGVLWTVYHGAGNKTLAGPKGPVTQMYTSAEGDLVGWPSP  1118
Qy  1029  CGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSG  1088
             |||| |.|||| |.||||||:..|:|:|::.||.:|.|.|.|.||:..|||||||:|..|
Db  1119  PGTKSLDPCTCGAVDLYLVTRNADVIPVRRKDDRRGALLSPRPLSTLKGSSGGPVLCPRG  1178
Qy  1089  HVIGMFTA---ARNSGGSVSQIRVRPLVCAGYHPQVTAHATLDTKPTVPNEYSVQILIAP  1145
             |:||:| |     |      :.|:|||||||||||.||  |.||.|:|.:|.|||: ||||
Db  1179  HAVGLFRAAVCARGVAKSIDFIPVESLDIAARTPSFSDNS---TPPAVPQTYQVGYLHAP  1235
Qy  1146  TGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKXMHATYGVNPNCYFNGKCTNTGASL  1205
             |||||||| | :.|.|:|:|||||||||:|:..|.|.|||||:||| :.| || ||| |
Db  1236  TGSGKSTKVPAYASQGYKVLVLNPSVAATLGFGAYMSKAHGINPNIRTGVRTVTTGDPI   1295
Qy  1206  TYSTYGMYLT-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATP  1263
             ||||||  || ||||:| ||||||||||:||||::||||.||::|..:|||||||||||
Db  1296  TYSTYGKFLADGGCSAGAYDVIICDECHSVDATTILGIGTVLDQAETAGARLVVLATATP  1355
Qy  1264  PGVIPTPHANITEIQLTDEGTIPFHGKIKEENLKKGRHLIPEATKGHCDELANELARKG   1323
             |||:.|||||||||||||:||||.||:||:.||:|||||:|.||:|.|||:|.||||:|
Db  1356  VNAVAYYRGLLDVSVIPTQGDVVVATDALMTGYTGDFDSVIDCNVAVTQVVDFSLDPTFT  1415
Qy  1324  ITAVSYYRGCCVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWGLS 1382
             : |:||||   ||||||||||||||||||||||||||||||||||||||||||||||||:
Db  1416  MGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWGLS   1442
Qy  1383  ITTQTVPQDAVSRSQRRGRTGRGRLGIYRYVSSGERPSGMFDSVVLCECYDAGAAWYELT  1442
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1476  ITTQTVPQDAVSRSQRRGRTGRGRLGIYRYVSSGERPSGMFDSVVLCECYDAGAAWYELT  1535
Qy  1443  STEAQTILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQL  1501
```

```
Db  1536  PAETTVRLRAYFNTPGLPVCQDHLEFWEAVFTGLTHIDAHFLSQTKQGGDNFAYLTAYQA  1595
Qy  1502  QLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADAC----------PGPEP------ 1545
             :  |                           ::|
Db  1596  TVCAR-----AKAPPSWD----------VMWK--------------CLTRLKPTLTGPTPLLYRLGAV 1635
Qy  1546  ------SEVTRY----QMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSIT   1594
Db  1636  TNEVTLTHPVTKYIATCMQADLEIMTS-TWVLAGGVLAAVAAYCLAT-------GCISII  1687
Qy  1595  S---VPTGATVAPVVDEEI-------VEECASFIPL-EAMVAAIDKLKSTI------TTT  1638
Db  1688  GRLHLNDQVVVAP--DKEVLYEAFDEMEECASKAALIEEGQRMAEMLKSKIQGLLQQATR  1745
Qy  1639  SPFTLETALE----KLNTFLGPHAATLAIIEYCCGLVTLPDNPFASCVFAFIAGITTPL   1694
             : :  :           :::      |: ||
Db  1746  QAQDIQPAIQSSWPKLEQFWAKHMWNFISGIQYLAGLSTLPGNPAVASMMAFSAALTSPL  1805
Qy  1695  PHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVG---FVFDMLGGGYAAA 1751
              ::| :   |:|| ||::|:|   |  :::| |  | :| | :|:|     :  :|| | |
Db  1806  PTSTTILLNIMGGWLASQIAPPAGATGFVVSGLVGAAVG---SIGLGKILVDVLAGYGAG  1862
Qy  1752  SSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTTAGPD--HWPNR  1809
             :|| | :|:::||| :: :|:    |:|:  | || || :||| :  :| ||||    ||:
Db  1863  ISGALVAFKIMSGEKPSVEDVVNLLPAILSPGALVVGVICAAILRRHVGQGEGAVQWMNR  1922
Qy  1810  LLTMLARSNTVCNEYFIATRDIRRKIIGILEASTPWSVISACIRWLHTPTEDDCGLI--   1866
             |: : |:.| | ||:| |: :|::  | :| |:|  ||:||||| : ::||    |:
Db  1923  LIAFASRGNHVAPTHYVAESDASQRVTQVLSSLT--ITSLLRRLHAWITEDCPVPCSG    1978
Qy  1867  AWGLEIWQYVCNPFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCG  1926
             :| :|||:||| :|| |::|| |  | || ||| || |||||:| |:..||: | |||||
Db  1979  SWLQDIWDWVCSILADFKNWLSA--KLLPKMFGLPFISCQKGYRGVWAGTGVWTTRCPCG  2036
Qy  1927  AELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARP-DPTDWTSLVVNYGVRDYC  1985
Db  2037  ANISGHVRMGTMKI-TGPKTCLNLWQGTPINCYTEGPCVPKPPPNYKTAIWRVAASEYV   2095
Qy  1986  KYEKMGDHIFVTAVSSPNVCFTQVP---PTLRAAVDGVQVQCYLGEPKTPWTTSACCY    2042
              : ::| |||  | ||||||  |     |:    |||||| |: || ||||: |  ||
Db  2096  EITQHGSFSYVTGLTSDNL---KVPCQVPAPEFFSWVDGVQVHRFAPIP---------   2141
Qy  2043  GPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSD---EAAVSALVFKQ  2099
Db  2142  GPFFRD---EVTFTVGLNSFVVGFQFPCDPEPDTEVLASMLTDPSHITAEAAARRLARGS  2198
Qy  2100  ELRRTNQLLEAISA---------------------GVDTTKLPA------------     2122
Db  2199  PPSQASSSASQLSAPSLKATCTTHKTAYDCDMVDANLFMGGDVTRIESDSKVIVVLSLDS  2258
Qy  2123  ----------PSI-BEVVVRKRQFRARTGSLTLPP--PPRSVPGVSCP--ESLQRSDPLEG 2168
              :   : ::
Db  2259  MTEVEDDREPSIPSEYLIRRKF----------PPALPWARPDYNPVIETWKR-------   2303
Qy  2169  PSNLPPS----------PPVLQLAMP--------------------MPLLGAGECNP   2195
Db  2304  PGYEPPTVLGCALPPTLQAPVPHRRRRAKVLTQDNVEGVLREMADKVFSPLQDNDSGH    2363
Qy  2196  FTAIGCAMTETGGGPDDLPSYPPKKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQ  2255
              : :                |:::   | ||: | ||:
Db  2364  STG------ADTGG---DSVQQPP-----DE--TAASEAGSLSSMPPLEGEPGDPDLE    2405
Qy  2256  SAPAKRPTKKKLGKSE---------------FSCSMSYTWTD-VISFKTASKVLSA   2295
              :  :    : :              || |||||:|| | ||::|:|:|
Db  2406  FEPA-GPAPPSEGECEVIDSDSKSWSTVSDQEDSVICCSMSYSWTGALITPCGPEEEKLP  2464
Qy  2296  TRAITSGFLKQRSLVVTEPRDAELRKQKVTINRQPLFPPSYHKQVRLAKEBAKSKVVGVM  2355
Db  2465  INPLSNSLMRFHNKVYSTTSRSASLRAKKVTFDRVQVLDAHYDSVLQDVKQAASKVSARL  2524
Qy  2356  WDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEAGEIPSHYR-QTVIVP  2414
             :  :|  | || ||:|     ||||:||    | ||||  || | ||::|   ||.:
Db  2525  LSVEEACALTPPHSAKSRY-GFGAKEVRS-LSRRAVNHIRSVWEDLLEDQHTPIDTTIMA  2582
```

```
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002318; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3033 AA;  330475 MW;  2OCC77C4F89C2AE3 CRC64;

Query Match             21.5%;
Best Local Similarity   29.9%;   Score 3303.5;  DB 2;  Length 3033;
                                 Pred. No. 1.7e-205;
Matches  985;  Conservative 443;  Mismatches 1163;  Indels 699;  Gaps 111;

Qy      8 TSPVPAPRTRKNKQTQASYPVSIK---------------TSVERGQRAKRKVQRDARPR    51
            |||:  ::  :                           ||:||||  | :|:|
Db      3 TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR    59

Qy     52 NYKIAGIHDGLQTLAQAALPAH------------------GWGRQDPRHKSRNL        87
               :    :                             ||||||||||||||
Db     60 GRRQPIPKDRRSTGKSWGKPGYPWPLYGNEGCGWAGWLLSPRGSRPTWGPTDPRHRSRNL   119

Qy     88 GILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLFV   144
          |  :|  |||| ||||  |||| ||    | | |   ::||||||:|||    |||| |
Db    120 GRVIDTITCGFADLMGYIPVVGAPV-GGVARALAHGVRVLEDGINYATGNLPGCSFSIFL   178

Qy    145 VCLLS-LACPCSGARVTDPDTNTTILNCCQRNQVIYCSPSTCLHEPGCVICADE-----   198
           ||||  ||| | |::: :|  ||  |||| ||:|   :||  ||:|||| |::
Db    179 LALLSCVTVPVGAVEVRN-ISSSYYATNDCSNNSITWQLTNAVLHLPGCVPCENDNGTLR   237

Qy    199 CWVPANPYISHPSNWTGTDSFLADHDFVMGALVTCDALDIGELCGACVLVGDWLV---   254
          ||  |:|  |  | |:| :  |::|:||| | ||   |:: : |||:|:| : ||
Db    238 CWIQVTPNVAVKHRGALTHN-LRTHVDMIVMAATVCSALYVGNVCGAVMLVSQALIVSPE   296

Qy    255 RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTKLASQVPYAIATMFS   314
          ||  :| |||:|| | ||||||||||:||||||||:|||||:| |:|||:|||| |::
Db    297 RH------NFTOECNCSIYQGQITGHRMAWDMMLNWSPTLTMILAYAARVPELVLEVVF   349

Qy    315 SVHYLAVGALIYYASRGKWYQLLALMLY-IEAT---SGNPI-------              352
           |  || ||| | ||:|||:|||| ||   ||    |||:
Db    350 GGHWGVMFGLAYFSMQGAWAKVIAILLLVAGVDATTYSSGTAVGSTLSSFTRLFTFGFQQ   409

Qy    353 ------------RVPTGCS------IAEF----CSPLMIPC-----                371
                       ||  ||        |||      ||||| |
Db    410 QLHLINTGSWHINRTALNCNDSLQTGFIASLFYANRPNSSGCPERLSSCRALDDFRIGW   469

Qy    372 --------------PCHSYLSENV-SEVICYSPKWTRPITLEYNN-------          401
                          ||  | |:||  ||| |  ||:|  ||: |:
Db    470 GALEYETNVTNDEDMRPYCWHYPPKPCGIVPARTVCGPVYCFTPS---PVVVGTTDRQGV   526

Qy    402 -SISWYPYTIPGARGCMVKFKNNT-----------WGC--------CRIR---NVP      434
            ||:|  |  ||| ||  |||::            ||         |||
Db    527 PTYSW-------GENETDVFLLNSTRPPRGAWFGCTWMNGTGFKTCGAPPCRIRRDFNST   580

Qy    435 SYCTMGTDAVWNDTRNTYEACGVTPWLT------------TAWHNGSALKLAILQ---YPGSK   482
           ||  |||| ||| |||||||||||||||
Db    581 LDLLCPTDCFRKHPDTTYIKCGAGPWLTPRCLIEYPYRLWHYPCTVNFTIFKVRMYVGGV   640
```

```
Qy   2415 KEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAVMGDAYGF-VDPRTR  2473
          |:|||  ||:|  ||| ||| ||:|: ||||||| :|:||:||| :| :|  ||   ||
Db   2583 KNEVFCVDPTKGGKKPARLIVYPDLGVRVCEKMALYDIAQKLPKAIMGPSYGFQYSPAER  2642

Qy   2474 VKRLLSMW--SPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQHRAGIHTIARQL  2531
          | |||  |    ||  | |||:||||||:| |||: | || |  || : | ||| ||||
Db   2643 VDFLLKAWRSKKDPMGFSYDTRCFDSTVTERDIRTEESIYQACSLPQEARVAIHSLTERL  2702

Qy   2532 YAGGPMIAYDGRRIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMKNPRFLICGDDC  2591
          | |||| :: || || ||||:||| || |:|::TC | |  |||:|| |:| ||:||||
Db   2703 YVGGPMTNSKGQSCGYRRCRASGVFTTSMGNTMTCYIKALAACKAAGIVDPVMLVCGDDL  2762

Qy   2592 TVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSNVTSGITKSGKP  2651
            |  |: | : : ::|  ||:|| ||| | | |  :|  |::||||||| ::  : || 
Db   2763 VVISESQONEEDERNLRAFTEAMTRYSAPPGDLPRPEYDLELITSCSSNVSVALDSRGRR  2822

Qy   2652 YYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHFEMEQMLFEDKLP  2711
          |||||||| |:  :     ||:| :  ||| :|  ||:| | : |||| | :|:| |::|
Db   2823 RYFLTRDPTTPITRAAWETVRHSPVNGWLGNIIQYAPTIWVRMVIMTHFFSILLAQDTLN  2882

Qy   2712 ETVTFPDWVGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSLTDMTMPPLRAW  2771
           :   :  |   |:| :::|| ||  :|| ||||:: :  :|| : ::::  |  ||||||
Db   2883 QNLNFEMYGAVYSVNPLDLPAIIERLHGLDAFSLHTYSPNELSRVAATLRKLGAPPLRAW  2942

Qy   2772 RKKARAVLASAKRRGGAHAKLARFLL-WHATSR----PLPDLDKTSVARYTTFNYCDVYS  2826
          : :|||| || |:||| |:   | :    | :      ||: |::::     :::    |
Db   2943 KSRARAVRASLIAQGGRAAICGRYLFNNAVKTKLKLTPLPEASRLDLSGWFT------VGA  2997

Qy   2827 PEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAI                         2863
           ||||| :    :  ||:  :||:|| |||:|:|
Db   2998 GGGDVFHSVSHARPRLL------LLCLLLSVGVGI                           3027
```

```
RESULT 7
Q7T7H9

ID   Q7T7H9        PRELIMINARY;     PRT;   3033 AA.
AC   Q7T7H9;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tanabe Y., Nagayama K., Enomoto N., Izumi N., Tazawa J., Sato C.,
RA   Watanabe M.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY232743; AAP55698.1; -.
DR   HSSP; Q8JYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
```

```
                                                        |||                    ||:|
Qy 1587 FAYLTAYQATVCAR----AKAPPPSWD-----              -VMWK-------CLTRLKPTLTGPT 1626
Db

Qy 1545 P--------------SEVTRY--------------      ---QMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGAT 1585
Db

Qy 1627 PLLYRLGAVTNEVTLTHPVTKYIATCMQADLEIMTS-TWVLAGGVLAAVAAYCLATGCVS 1685
Db

Qy 1586 CVRRCWSITSVPTGATVAPVVDEEI------   VEECASFIPL-EAMVAAIDKLKSTI--- 1635
Db

Qy 1686 IIGRIHLNDQV----VVAP--DKEVLYEAFDEMEECSSKAALIEEGQRIAEMLKSKIQGL 1739
Db

Qy 1636 ---TTTSPFTLETALE----KLNTFLGPHAATILAIIEYCCGLVTLPDNPFASCVFAFIA 1688
Db

Qy 1740 LQQATRQAQDIQPAIQSSWPKLEQFWAKHMWNFISGIQYLAGLSTLPGNPAVASMMAFSA 1799
Db

Qy 1689 GITTPLPHKIKMFLSLFGGAIASKLTDARGALAFWMAGAAGTALGTWTSVG---FVFDML 1745
Db

Qy 1800 ALTSPLSTSTTILLNIMGGWLASQIAPPAGATGFVVSGLVGAAVG--SIGLGKILVDIL 1856
Db

Qy 1746 GGYAAASSTACLTPKCLMGEWPTMDQLAGLVYSAFNPAAGVGVLSACAMFALTTAGPD- 1804
Db

Qy 1857 AGYGAGISGALVAFKIMSGEKPSVEDVVNLLPAILSPGALVVGVICAAILRHVGQGEGA 1916
Db

Qy 1805 -HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTEDDC 1863
Db

Qy 1917 VQWMNRLIIAFASRGNHVAPTHYVAESDASQRVMQMLSSLT---ITSLLRLRHTWITEDC 1972
Db

Qy 1864 GLI---AWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ 1920
Db

Qy 1973 PVPCSGSWLRDVWDWVCSILIDFKNWLSS--KLLPKMPGLPFISCQKGYRGVWAGTGVMT 2030
Db

Qy 1921 ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARP-DPTDWTSLVVNY 1979
Db

Qy 2031 TRCPCGANISGHVRMGTMKI-TGPKTCLNLWQGTFPINCYTEGPCVPKPPNVKTAIWRV 2089
Db

Qy 1980 GVRDYCYEKMGDHIFVTAVSSPNVCFTQVP---PTLRAAVADGVQVQCYLGEPKTPWT 2036
Db

Qy 2090 AASEYIEVTQHGSFSYVTGLTSDNL---KVPCQVPAPEFFSWNDGVQIHRFAPIP------ 2141
Db

Qy 2037 TSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSD---EAAVS 2093
Db

Qy 2142 ------GPFFRD---EVTFTVGLNSLLVVGSQLPCDPEPDTEVLASMLTDPSHITAEAAAR 2192
Db

Qy 2094 ALVFKQELRRTNQLLEAISA--   GVDTTKLPA------- 2122
Db

Qy 2193 RLARGSPPSQASSASQLSAPSLKATCTTHRMAYDCDMVDANLFMGGDVTRIESDSKVIV 2252
Db

Qy 2123 ------PSI-EEVVVRKRQFRARTGSLTLPP--PPRSVPGVSCP--ESLQR 2162
Db

Qy 2253 LDSLDTMTEVEDDREPSVPSEYLVRKRKF---PPALPPWARPDYNPPVIETWKR 2303
Db

Qy 2163 SDPLEGPSNLPPSPPVLQAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKKE- 2221
Db

Qy 2304 ------PGYEPP----   TVLGCALPPTPQAP--VP--PPRRRR 2331
Db

Qy 2222 ----   VSEWSDESWSTA 2244
Db

Qy 2332 AKVLTQDNVGEVLREMADKVLSPLQDHNDSGHSTGADTGGDSVQPSDETAASEAGSSSS 2391
Db

Qy 2245 YPKIRGK--------DSTQSAPAKR----   PTKKKLGKSEFSCCSMSYYTWTD- 2282
Db

Qy 2392 MPPLEGEBPGDPDLEFEPAESAPPSEGECEVIDSDSKSWATVSDQEDSVICCSMSYSWTGA 2451
Db

Qy 2283 VISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQVR 2342
Db

Qy 2452 LITPCGPEEEKLPINPLSNSLMRFHNKVYSTTSRSASLRAKKVTFDRVQVLDAHYDSVLQ 2511
Db

Qy 2343 LAKEASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEAGE 2402
Db

Qy 2512 DVKRAASKVSARLLSIEBACALTPPHSAKSRY-GFGAKEVRS-LSRRAVNHIRSVWEDLL 2569
Db

Qy 2403 IPSHYR-QTVIVPKEBVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYGQVAPDVVKAVM 2461
Db
```

```
Qy 483 EMFKPHNWMSGHLYF--------EGSDTPIVYFYDPVNSTLLPPERWARLPGTPPV 530
Db 641 E------HRMSAACNFTRGDRCRLEDRDRGQQSPLLH------ST---TEWAVLPCS--- 682

Qy 531 VRGSWLQVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGL 590
Db 683 -----FSDLPALSTGLMHLHQNIVDVQVLYGLSPAVTRYIVKWEWVVLLFLLL 730

Qy 591 CGSKYLILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIPFICCY 650
Db 731 ADAR---VCACLWMLI-------ILGQAEAALEKLII---- 757

Qy 651 LRCRLRYAALLGFVPMAAGLPLTFFVAAAAAOPDYDWVRLLVAGLVLWAGRNRGHRIAL 710
Db 758 ------LHSASASANGPLWFFIFFTTA-------WYLKGRVVPVATYS---- 793

Qy 711 LVGPWPLVALLTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGF---FAHLLPRCAL 767
Db 794 VLGLWSF--LLLVLALPQQAVALDTAEQGELGL--VILVIISIFTLTPAYKILLSRSVW 848

Qy 768 VNSYLW------QRW------ENWFW-NVTLRPERFFLV---LVCFPGATYDALV 806
Db 849 WLSYMLVLAEAQIQQWVPPLEARGGRDGIIWAAVILHPRLVFEVTKWLLAILGSAY---- 904

Qy 807 TFCVCHVALLCLTSSAASFPGTDSRVRAHRM-----LVRLGKCHAWYSHYVLKFFLLVFG 861
Db 905 ----LLKASLLRVPYF-----VRAHALLRVCTLVR----HLAGAKYI-QMLLITLG 946

Qy 862 E-NGVFFYKHLHGDVLPNDFASKLPLQ----   EP--FFPPEGKARVVRNEG 904
Db 947 RWTGTYIYDHLS------   PLSTWAAQGLRDLAVAVEPVVFSPMEKKVIVWGAE- 993

Qy 905 RRLACGDTVDGLPVVARLGDLVFAGLAMPPD-----GWAITAPFTLQCLSERGTLSAMAV 959
Db 994 -TVACGDILHGLPVSARLGKEVLLG---PADSYTSKGWKLLAPITAYTQQTRGLLGAIVV 1049

Qy 960 VMTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAA 1019
Db 1050 SLTGRDKNEQAGQVQLLSSVTQSFLGTSISGVLWTVYHGAGNKTLAGPKGPVTQMYTSAE 1109

Qy 1020 NDQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSS 1079
Db 1110 GDLVGWPSPPGTKSLDPCTCGAVDLYLVTRNADVIPVRRKDDRRGALLSPRPLSTLKGSS 1169

Qy 1080 GAPILCSSGHVIGMFTA---ARNSGGSVSQIRVRPLVCAGVHPQYTAHATLDTKPTVPNE 1136
Db 1170 GGPVLCPRGHAVGLFRAAVCARGVAKSIDEFVPVESLAIATRTPSFSDNS---APPAVPQT 1226

Qy 1137 YSVQILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNG 1196
Db 1227 YQVGYLHAPTGSGKSTKVPAAYASQGYKVLVLNPSVAATLGFGAIMSKARGINPNIRTGV 1286

Qy 1197 KCTNTGASLTYSTYGMYL-TGACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVR 1254
Db 1287 RTVTTGDPITYSTYGKFIANGGCSAGAYDVIICDECHSSVDATTILGIGTVLDQAETAGAR 1346

Qy 1255 LVVLATATPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKGHCDE 1314
Db 1347 LVVLATATPGTVTTPHSNIEEVALGHEGEIPFYGKAIPLAFIRGGRHLIFCHSKKKCCDE 1406

Qy 1315 LANELARKGITAVSYYRGCDISKIP-EGDCVVATDALCTGYTGDFDSVVDCSLMVEGTC 1373
Db 1407 LAAALRGMGVNAVAYYRGLDVSVIPTQGDVVVVATDALMTGYTGDFDSVIDCNVAVTQIV 1466

Qy 1374 HVDLDPTFTMGVRVCCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFD 1433
Db 1467 DFSLDPTFTITTQTVPQDAVSRSQRRGRTGRGRLGIYRYVSSGERPSGMFDSVVLCECYD 1526

Qy 1434 AAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADN 1492
Db 1527 AGAAWYELTPAETTVRLRAYFNTPGLPVCQDHLEFWBAVFTGLTHIDAHFLSQTKQGGDN 1586

Qy 1493 YVLLTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADAC-------PGPB 1544
Db
```

```
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_Pol_DS_PS
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3033 AA;  329974 MW;  9448C41AFC5BB89E CRC64;

Query Match             21.5%;  Score 3303.5;  DB 2;  Length 3033;
Best Local Similarity   30.2%;  Pred. No. 1.7e-205;
Matches  993;  Conservative  434;  Mismatches 1163;  Indels  701;  Gaps  110;

Qy     8 TSPVAPRTRKNKQTQASYPVSIK------------------TSVERGQRAKRKVQRDARPR 51
           ::: :  :: ||:  | :                        ||    | :  : : ||
Db     3 TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR 59

Qy    52 NYKIAGIHDGLQTLAQALPAH------GWGRQDPRHKSRNL 87
          : :                  :      || || :||:|||
Db    60 GRRQPIPKDRRSTGKSWGKPGYPWPLYGNEGCGWAGWLLSPRGSRPTWGPTDPRHRSRNL 119

Qy    88 GILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLFV 144
          | ::| : ||:||:||:| |  |: :  |  : :|:|| ||||:|:|||   ||  : ||
Db   120 GRVIDTITCGFADLMGYIPVVGAPV-GGVARALAHGVRVLEDGINYATGNLPGCSFSIFL 178

Qy   145 VCLLS-LACPCSGARVTDPDTNTTILNCCQRNQVIYCSPSTCLHEPGCVICADE------ 198
          : ||| : |  | | :    :   :    | :  ::    : : ||||| | :
Db   179 LALLSCVTVPVSAVEVRN-ISSSYYATNDCSNNSITWQLANAVLHLPGCVPCENDNGTLR 237

Qy   199 CWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV--- 254
          ||  :  ||   | : |   |  | ||::|| ||| || |  :  |  |::  : | :
Db   238 CWIQVTPNVAVKHRGALTHN-LRTHVDMIVMAATVCSALYVGDVGVGAVMIVSQALIVSPE 296

Qy   255 RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKGVEAVIPLTKLASQVPYAIATMFS 314
          ||         |    |: | |   | :   :  : || || | ::  |  |:: || :
Db   297 RH------NFTOECNCSIYQGHITGHRMAWDMLMNWSPTLTMILAYAARVPELVLEVVF 349

Qy   315 SVHYLAVGALIYYASRGKWYQLLLALMLY--IEAT---------- 347
          |  : |:| |: : |:| : |::| ||:|  |:||
Db   350 GGHWGVVFGLAYFSMQGAWAKVIAILLLVAGVDATTYSTGAQAGRAASGIANLFTPGAKQ 409

Qy   348 ----------------------------SGNPIRVPT------GC 358
                                      |||:||||
Db   410 NIQLINTGSWHINRTALNCNDSLQTGFLASLFYTKSFNSSGCPERLSSCRGLDDFRIGW 469

Qy   359 STAEFCSPL----------MIPCPCHSYLSENV-SEVICYSPKWTRPITL- 397
          ||  ||| |          :||| | | | | : || :| || | |  |
Db   470 GTLEYETKVTNDEDMRPYCWHYPPKPCGIVSAKTVCGPVYCFTPS---PIVVGTDKQGV 526

Qy   398 -EYN------NSI------SWYPYTIPGARGCMVKFKNNTWGC--CRIR----NVPS 435
           || |      |||      |:|| |:|| |  || | | |||  ||||
Db   527 PTYNWGDNETDVFLLNSTRPPQGAWFGCTWMNGTGF----TKTCGAPPCRIRRDFNSTL 581

Qy   436 YCTMGTGTDAVWNDTRNTYEACGVTPWLT----------TAWHNGSALKLAILQ--YPGSKE 483
          :|   | |:  :    || || ||| |||          || | | || ::|  |||| |
Db   582 DLLCPTDCFRKHPDATYIKCGAGPWLTPRCLVEYPYRLWHYPCTVNFIFKVRMYVGGVE 641

Qy   484 MFKPHNWMSGHLYF--------EGSDTPIVYFYDPVNSTLLPPERWARLPGTPVV 531
          : ||  |  | |:         |||:|| | ||| |||:|| |   | |||| ||
Db   642 -----HRLSAACNFTRGDRCRLEBDRDRGQQSPLLH------ST----TEWAVLPCS- 682
```

```
Db  2570 EDQHTPIDTTVMAKNEVFCVDPAKGGKKPARLIVYPDLGVRVCEKMALYDIAQKLPKAIM 2629
          ::  :  :| ::  | :  ||      :|||||:||||| :   : ||||| | :|:
Qy  2462 GDAYGF-VDPRTRVKRLLSMW--SPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSD 2518
          | : || ||| |  |:||  :    : |  |  |:||||:||:|||:| |||||:| |
Db  2630 GPSYGFQYSPAERVDFLLKAWRSKKDPMGFSYDTRCFDSTVTERDIRTEESIYQACSLPE 2689
Qy  2519 QHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGGVVTTSSSNSLTCWLKVNAAAEQAG 2578
          :  |||:::  || ||||| : : | |||||| |||||||| |||| ||::| || |:||
Db  2690 EARTVIHSLTERLYVGGPMTNSKGQSCGYRRCRASGVFTTSMGNTMTCYIKALAACKAAG 2749
Qy  2579 MKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCCS 2638
          : :|  |:|||| :||  | | :: :::|| | :|| |:|:|:|| | | |: |:|||||
Db  2750 IMDPMMLVCGDDLVVISESQGNEEDERNLRAFTEAMTRYSAPPGDLPRPEYDLELITSCCS 2809
Qy  2639 SNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAV 2698
          |||: | | | : :|||||| : | :  :  : | ||  | |||::::|| ||   :::
Db  2810 SNVSVALDSRGRRRYFLTRDPTTPITRAAWETVRHSPVNSWLGNIIQYAPTIWVRMVIMT 2869
Qy  2699 HFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQ 2758
          || : :| | :|: :  |: | :  :| ::||  ||  : |::||| :  ::| | ||:|
Db  2870 HFFSILLAQDTLNQNLNFEMYGAVYSVNPLDLDPAIIERLHGLDAFSLHTYSPHELSRVAA 2929
Qy  2759 SLTDMTMPPLRAWRKKARAVLASAKRGGAHAKLARFLL-WHATSR----PLPDLDKTSV 2813
          :| ::  ||||| : ||| : ::||  |:|  |   |: |:| ||    ||| |::
Db  2930 TLRKLGAPPLRAWKSRARAVRASLIAQGSRAAICGRYLFNWAVKTKLKLTPLPEASRLDL 2989
Qy  2814 ARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKFLVKYLAVIVFALGLIAVGLAI 2863
          : :  | :  :   : |:    : || :  : ::||  : |:::  | ||:|:|
Db  2990 SGWFT-----VGAGGGDIYHSVSHARPRLL------LLCLLLSVGVGI 3027
```

RESULT 8
Q7T714
```
ID   Q7T714      PRELIMINARY;      PRT;   3033 AA.
AC   Q7T714;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tanabe Y., Nagayama K., Enomoto N., Izumi N., Tazawa J., Sato C.,
RA   Watanabe M.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY232738; AAP55693.1; -.
DR   HSSP; Q8JYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
```

us-10-009-002-2.rup

```
Qy  1546 ------------SEVTRY----QMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWS 1592
                                  ::  :  |  |:  ::  :|::   :::
Db  1634 AVTNEVTLTHPVTKYIATCMQADLEVMTS-SWVLAGGVLAAVAAYCLAT------GCIS 1685

Qy  1593 ITS----VPTGATVAPVVDEE------EIVEECASFIPL-EAMVAAIDKLKST-----T 1636
          ||         |        ::       |||:
Db  1686 IIGRLHLNNQVVVAP--DKEILYEAFDEMEEECASKAALIEEGQRMAEMLKSKIQGLLQQA 1743

Qy  1637 TTSPFTLETALE-----KLNTFLGPHAATILAIIEYCCGLVTLPDNPFASCVFAFIAGITT 1692
                :          :| |:   |:  ::  ::
Db  1744 TKQAQDIQPAVQSSWPKLEQFWAKHMWNFISGIQYLAGLSTLPGNPAVASMMAPSAALTS 1803

Qy  1693 PLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAGTALGTWTSVG--FVFDMLGGYA 1749
          ||| | :|  :: ||  :|:|  :  | :: :||:|  |   :       :  :|||
Db  1804 PLPTSTTLLNIMGGWLASQIAPPAGATGFVVSGLVGAAVG---SIGLGKILVDVLAGYG 1860

Qy  1750 AASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTTAGPD--HWP 1807
           | :|| | ||:|:||:||:: || ::| |  | |:|| |::||:: | |       :
Db  1861 AGISGALVAFKIMSGEKPTVEDAVNLLPSILSPGALVVGVICAAILRRHVGQGEGAVQWM 1920

Qy  1808 NRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTEDDCGLI- 1866
          ||| | :||| |   :   ||    |:|| ||:|| :|: :::  | | | |:|  |
Db  1921 NRLIAFASRGNHVAPTHYVTESDASQRVIQVLSSLT----ITSLLRRLHAWITEDCPVPC 1976

Qy  1867 --AWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQARCP 1924
             :| |  |: | |::|  :  |:| : |   |||| :|||||:|  || |: ::|||
Db  1977 SGGSWLRDIWDWVCSILTDFKNWLSS--KLLPKMPGLPFISCQKGYRGWAGTGVMTTRCP 2034

Qy  1925 CGAELIFSVENGFAKLYKGPRTCSNYNRGAVPVNARLCGSARP-DPTDWTSLVVNYGVRD 1983
          |||  :   |    :| ||||||:|:|| |: || |||   |    |::|  |||:|| |
Db  2035 CGANISGHVRMGTMKI-TGPKTCLNLWQGTFPINCYTEGPCVKPPPNYKTAIWRVAASE 2093

Qy  1984 YCKYEKMGDHIFVTAVSSPNVCFTQVP---PTLRAAVADVGVQVCYLGEPKTPWTTSAC 2040
           | |  :|  : ::| ||  ||| : :    |   | |:| || ||: : :  | :|
Db  2094 YVEVTQHGSFSYITGLTSDNL---KVPCQVPAPEFFSWVDGVQIHRFAPIP------- 2141

Qy  2041 CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSD--EAAVSALVF 2097
            |  |  :|                  |: | | |:|| :    | |    || :|
Db  2142 --GPFFRD---EVTFTVGLNSFGVGSGQLPCDPEPDTEVLASMLTDPSHITAEAAARRLAR 2196

Qy  2098 ----KQELRRTNQL-----LEA------------ISAGVDTTKLPA 2122
            || : :  :|       :| |            :| |:|| ||
Db  2197 GSPPSQASSSASQLSAPSLKATCTTHKMAYDCEMVDANLSRGGDVTRIESDSKIVLDSL 2256

Qy  2123 -----------PSI-EEVVVRKRQFRARTGSLTLPP--PPRSVPGVSCP--ESLQRSDPL 2166
                      ||  | : :|::: :  |  :  ||   || :
Db  2257 DSMTEVDDDREPSVPSEYLIRRKKF----PPALPPWARPDYNPPTIEIWKR----- 2303

Qy  2167 EGPSNLPPS-------------PPVLQLAMPMP------------ 2186
           | : :||:              |   :  |
Db  2304 --PGYEPPTVLGCALPPSLQAPVPPPRRRRRAKVLTQDNVEGVLREMADKVLSPLQDHNDS 2361

Qy  2187 --LLGA---GECNPFTAIGCAMTETG-----------GGPD----DLPSYPPKK----EV 2222
             :||     |  |    |  | ||           ||        | :||||     |
Db  2362 GHSTGADTGGDSAQQPSDETAASETGSLSSMPPLEGEPGDPDLEFEPAGSAPPSEGECEV 2421

Qy  2223 SEWSDESWSTATTASSVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCCSMSYTWTD 2282
                                                            || ||||| |
Db  2422 --------------------------------------------ICCCSMSYSWTG 2450

Qy  2283 -VISFKTASKVLSATRAITSGFLKQRSLVYYTEPRDAELRKQKVTINRQPLFPPSYHKQV 2341
           :                  || :   |::   |||  |||:|:||     |  ::: :|
Db  2451 ALITPCGPEBEKLPINPLSNSLMRFHNKVYSTTSRSASLRAKKVTFDRVQVLDAHYDSVL 2510

Qy  2342 RLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEAG 2401
           : :: ||||: |    :|| | ||||:|||     || ||| |:| |||:|: |||| |
Db  2511 QDVKQAASKVSARLLSVBEACALTPPHSAKSRY-GFGAKEVRS-LSRRAVNHIRSVWEDL 2568

Qy  2402 EIPSHYR-QTVIVPKEEVFVKTPQKPKQKPPRRLISYPHLEMRCVEKMYYGQVAPDVVKAV 2460
           :      ||: :  ::|| : ||:    |||::|:||  : ||||||  | |||||:|:|
Db  2569 LEDQHTPIDTTIMAKNEVFCVDPTKGGKPKARLIVYPDLGVRVCEKMALYDIAQKLPQAI 2628

Qy  2461 MGDAYGF-VDPRTRVKRLLSMW--SPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLS 2517
```

```
Qy   532 RGSWLQVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVLILLGLC 591
           :|::         :|:  ||  : :   :|:|||: |   ||| |||  ||| | |||:
Db   683 ------------FSDLPALSTGLLHLHQNIVDVQYLYGLTPAITRYIVKWEWVVLLFLLLA 731

Qy   592 GSKYLIAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYL 651
           :::         |  |:      ::            ::|                 |
Db   732 DAR--VCACLWMLI------------------IILGQAEAALEKLII------- 757

Qy   652 RCRLRYAALLGFVPMAAGLPLTFVAAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALL 711
                        |  |:  :    | |:    |||:       ::  |  : :
Db   758 ------LHSASAAASANGPMWFFIFFIAA----WYLKGRVVPMATYS-------V 794

Qy   712 VGPWPLVALLTLLHLVTPASAFDTEIIG--GLTIPPVVALVVMSRFGFFAHLLPRCALVN 769
           :| | ||||||:|:|:||||| :  |:   ||  : |  :|:     ||| ||    |:
Db   795 LGLWSF-LLLVLALPQQAYALDTAEQGELGLVILAIISIFTLT--PAYKILLSRSVWWL 850

Qy   770 SYLW-----QRW-------ENMFW-NVTLRPERFFLV----LVCFPGATYDALVTF 808
           |:             :|  :  :|   :  |||  ::    |  | |||  |:
Db   851 SYMLVLAEAQIQQWVPPLEARGGRDGIIWVAVILHPRLVFEITKWLLAILGPAY------ 904

Qy   809 CVCHVALLCLTSSAASFFGTDSRVAHRMLVRLGKCHAWYSHYVLKFFLLVFGE-NGVFF 867
                ||  |: :  :       :|| ||||  :|   ::  |:   : : ||
Db   905 -----LLRASLLRIPYF-----VRAHALLRVCTLVRRIAGAKYIQVLLLTIGRWTGTYI 953

Qy   868 YKHLHGDVLPNDFASKLPLQ----EP---FFPFEGKARVYRNEGRRLACGD 911
           | |||                   |:    | |  ||::: :  |::||||
Db   954 YDHLS-----------PLSTWAAQGLRDLAVAVEPVVFSPMEKRVIVWGAE--TVACGD 999

Qy   912 TVDGLPVVARLGDLVFAGLAMPPD-----GWAITAPPTLQCLSERGTLSAMAVVWTGIDP 966
           ::  ||| ||:|: : || | |       |: : || : |  |  |:| |:|| | |   |
Db  1000 ILHGLPVSARIGREVLLG----PADSYTSKGWKLLAPITAYTQQTRGLLGAIVVSLTGRDK 1056

Qy   967 RTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDIYQ 1026
              :|   :|:|| || :|  ::     | ||| :   |  |   :  ::   | :  :
Db  1057 NEQAGQVQVLSSVTQSFLGTSISGVLWTVVHGAGNKTLAGPKGPVTQMYTSAEGDLVGWP 1116

Qy  1027 PPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPWCVCGALPMAVAKGSSGAPILCS 1086
           ||      ||| |||   ||||| : :  ::     :||:|||| | ||| ||||:|||
Db  1117 SPPGTKSLDPCTCGAVDLYLVTRNADVIPVRRKDDRRGALLSPRPLSTLKGSSGGPVLCS 1176

Qy  1087 SGHVIGMFTAA---RNSGGVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILI 1143
            |||:|:| ||   |:  |  ::      | | | ||| || ||           | :|
Db  1177 RGHAVGLFRAAVCTRGVAKSIDFIPVESLDIAARTPSFSDNS---TPPAVPQSYQVGYLH 1233

Qy  1144 APTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGA 1203
           ||||||||||:|  || :: | ||||||||: :   ||: || |||| : | |  ||
Db  1234 APTGSGKSTKVPAAYASQGYKVLVLNPSVAATLGFGAYMSKAHGINPNIRTGVRTVTTGD 1293

Qy  1204 SLTYSTYGMYLT-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATA 1261
            :|||||| :|  | |   ||||||||||:| |||| |||| ||  |   :|||||||||
Db  1294 PITYSTYGKFLADGGCSAGAYDVIICDECHSVDATTLGIGTVLDQAETAGARLVVLATA 1353

Qy  1262 TPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELAR 1321
           ||||  ||||:||  :    :  ||:||| | : |  |||||||    |||||| ::| |
Db  1354 TPPGTVTTPHSNIEEVALGHSEGBEIPFYGKAIPLAFIKGGRHLIFCHSKKKKCDELASALRG 1413

Qy  1322 KGITAVSYYRGCDISKIP-EGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT 1380
            |: ||:|||| |:| :|  ||| |||||||:|||||||||| ||:: :       |||
Db  1414 MGVNAVAYYRGLDVSVIPTQGDVVVVATDALMTGYTGDFDSVIDCNVAVTQIVDFSLDPT 1473

Qy  1381 FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVRAFDAAKAWYG 1440
           || : :   | |:::||||||||||| ||| ||  ::|||| : || :  | ||| ||
Db  1474 FTITTQVTVPQDAVSRSQRRGRTGRGRLGIYRYVSSGERPSGMFDSVVLCECYDAGAAWYE 1533

Qy  1441 LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAA 1499
           |   |  |   || |:|||| :   |  || :    ||| ||   : :::| | ||||
Db  1534 LTPAETTVRLRAYFNTPGLPVCQDHLEFWEAVFTGLTHIDAHFLSQTKQGGDNLAYLTAY 1593

Qy  1500 QLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADAC------PGPEP------ 1545
           |     ::|:   :   :| |:| ||||| | :                ||
Db  1594 QATVCAR----AKAPPSWD------VMWK------CLTRLKPTLTGPTPLLYRLG 1633
```

```
Qy  2629 MGPSYGFQYSPAERVDFLLKAWGSKKDPMGFSYDTRCFDSTVTERDIRTEESSIYQACSLP 2688
         ||:|||  |    |           ||   || ||      || |  |::    |  |
Db  2518 DQHRAGIHTIARQLYAGGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQA 2577

Qy  2689 QEARTVIHSLTERLYVGGPMTNSKGQSCGYRRCRASGVFTTSMGNTMTCYIKALAACKAA 2748
         |||||||  ||   ||||||||||| |||||||| ||||||||||||||| | | ||
Db  2578 GMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSC 2637

Qy  2749 GIVDPVMLVCGDDLVVLSESQGNEEDERNLRAFTEAMTRYSAPPGDLPRPEYDLELITSC 2808
         ||| || |||||| :| | :  |::   ||| ||   ||:| |:     ::  : :||
Db  2638 SSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLA 2697

Qy  2809 SSNVSVALDSRGRRRYFLTRDPTTPITRAAWETVRHSPVNSWLGNIIQYAPTIWVRMVIM 2868
         ::|:: :    |  :|::|   |  |  :|    |: ||   |  :: | ||: |:  :|
Db  2698 VHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVS 2757

Qy  2869 THFFSILLAQDTLNQNLNFEMYGAVYSVNPLDLPAIIERLHGLEAFSLHTYSPHELSRVA 2928
         |||  : |   |:|::|||  | : |||  || || ||| :||||||:|:|:||:||:||
Db  2758 QSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WHATSR-----PLPDLDKTS 2812

Qy  2929 ATLRKLGAPPLRAWKSRARAVRASLIAQGGRAAICGRYLFNWAVKTKLKLITPLPEARRLD 2988
         | || ||||||||| |||||:|||||| |||||||:  ||: |     |  |||| ||||
Db  2813 VARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAI 2863

Qy  2989 LSGWFT-----VGAGGGDIFHSVSHARPRLL-------LLCLLLSVGVGI 3027
```

RESULT 9

Q9IZA1

```
Q9IZA1;
ID   Q9IZA1          PRELIMINARY;      PRT;   3033 AA.
AC   Q9IZA1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MD2b-1;
RA   Itakura J., Nagayama K., Enomoto N., Kurosaki M., Watanabe H.,
RA   Sato C.;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF238486; AAF59945.1; -.
DR   HSSP; Q8JYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC heme_BS.
DR   InterPro; IPR001531; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3033 AA;   330724 MW;   FE04FEC7C385A13A CRC64;

Query Match             21.5%;     Score 3303.5;  DB 2;   Length 3033;
Best Local Similarity   30.3%;     Pred. No. 1.7e-205;
Matches  991;  Conservative 447;  Mismatches 1176;  Indels 653;  Gaps 112;

Qy      8 TSPVPAPRTRKNKQTQASYPVSIK------------------------TSVERGQRAKRKVQRDARPR 51
          :|:||  :||:                                      ::||
Db      3 TNPKP---QRKTKRSTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR 59

Qy     52 NYKIAGIHDGLQTLAQAALPAH------------------------GWGRQDPRHKSRNL 87
          :                                            ||||||||:|
Db     60 GRRQPIPKDRRSTGKSWGKPGYPWPLYGNEGCGWAGWLLSPRGSRPHWGPTDPRHRSRNL 119

Qy     88 GILLDYPLGWIGDVTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG--WFGVHLFV 144
          ||:  |||:||:| ::|:|||  ||||||||:|  |:||||||:||||  ||::||:|
Db    120 GRVIDTITCGFADLMGYIPVVGAPV-GGVARALAHGVRVLEDGINYATGNLPGCSFSIFL 178

Qy    145 VCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVICADE------- 198
          |:||   | | |::|    :::  :|    |  ||   : |  :|||  ||:|
Db    179 LALLSCVTPVPSAVEVRN-ISSSYYATNDCSNSSITWQLTNAVLHLPGGVPCENDNGTLR 237

Qy    199 CWVPANPVYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV---- 254
          |  | : :|:|  :|:||||| | | |:|  ||:| | ||||| |  |:|:||::|
Db    238 CWIQVTPNVAVKHRGALTHN-LRTHVDMIVMAATVCSALYVGDVCGAVMIVSQALIVSPE 296

Qy    255 RIHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTKLASQVPYAIATMFS 314
          |:||   ||                                  ::::||  |||:||:|
Db    297 RH-----NFTQECNCSIYQGHITGHRMAWDMLNWSPTLTMILAYAARVPELALEIVF 349

Qy    315 SVHYLAVGALIYYASRGKWYQLLLALMLY--IEATSGNPIRVPTGCSIAEFCSPL----- 367
          |  :||:|   || | | |:|: |:||:   : :|  :|:  :|::|||
Db    350 GGHWGVAFGLAYFSMQGAWAKVIALLLVAGVDATT---YSTGAQVGQTVSGFAGMFR 404

Qy    368 ------------------MIPCPCHSYLSENVSEVICYSPKWT------------RPI--- 395
                            :  | | | : :|   | || |:           |::
Db    405 SGSRQNIQLINTNGSWHINRTALNCNDSLQTGFMASLFYAKRFNSSGCPERLSSCRRLDD 464

Qy    396 ----------------TLEYNNSIS---------------------WYPYT------- 409
                          ::|:|:|:                     ||| |
Db    465 FRIGWGTLEYETNVTNDEDMRPYCWHYPPKPCGIVSDKTVCGPFYCYTPSPVVVGTSDKQ 524

Qy    410 -IP----GARGCMVKFKNNT----------------------CRIR------NVPSY 436
             |    ||||| | |::|                       |:|:      || |
Db    525 GVPTYSWGENETDVFLINSTRPPRGDRFGCTWMDETGFTKTCGAPPCRIRRDYNSSIDLL 584

Qy    437 CTMGTDAVWNDTRNTYEACGVTPWLT------------TAWHNGSALKLAILQ---YPGSKEM 484
              :| |::|  :::||:|  |||:|            ||  ||||:|  ::|   || :|:
Db    585 CP---TDCFRKHPGATYIKCGAGPWLTPRCVVDPYPYRLWHYPCNVNFTIFKVRMVVGGVE- 641

Qy    485 FKPHNWMSGHLYF----------------------EGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVR 532
            | |::  |:|:                       :|:| |::   ::|     |  |:   |  ||| |
Db    642 ----HRLTAACYFTPGRDRCQLEDRDRGQQSPLLH-----ST----TEWAVLPCT- 682

Qy    533 GSWLQVPQGFYSDVKDLATGLIITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCG 592
```

```
Db 1630 YRLGAVTNEVTLTHPVTKYIATCMQADLEIMTS-SWVLAGGVLAAVAYCLAT------ 1681
Qy 1589 RCWSITS---VPTGATVAPVVDEE------EIVEECASFIPL-EAMVAAIDKLKSTI--- 1635
Db 1682 GCISIIGRLHLNDQVVVAP--DKEILYEAFDEMEECASKAALIEEQRMAEMLKSKIQGL 1739
Qy 1636 ---TTTSPFTLETALE----KLNTFLGPHAATILAIIEYCCGLVTLPDNPFASCVFAFIA 1688
Db 1740 LQQATRQAQDIHPAIQSSWPKLEQFWAKHMWNFISGIQYLAGLSTLPGNPAVASMMAFSA 1799
Qy 1689 GITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMWAGAAGTALGTWTSVG---FVFDML 1745
Db 1800 ALTSPLPTSTTILLNIMGGWLASQIAPPAGATGFVVSGLVGAAVG---SIGLGKILVDVL 1856
Qy 1746 GGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTTAGPD- 1804
Db 1857 AGYGAGISGALVAFKIMSGEKPSVEDVVNLLPAILSPGALVVGVICAAILRRHVGQGEGA 1916
Qy 1805 -HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTEDDC 1863
Db 1917 VQWMNRLIAFASRGNHVAPTHYVAESDASQRVTQLLGSLT----ITSLLRRLHAWITEDC 1972
Qy 1864 GLI---AWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ 1920
Db 1973 PVPCSGSWLRDVWDWVCSILITDFKNWLSS--KLLPKMPGLPFISCQKGYRGVWAGTGVMT 2030
Qy 1921 ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARP-DPTDWTSLVVNY 1979
Db 2031 TRCPCGANISGHVRMGTMKI-TGPKTCLNLWQGTFPINCYTEGPCVPKPPPNYKTAIWRV 2089
Qy 1980 GVRDYCKYEKMGDHIFVTAVSSPNVCFTQVP--PTLRAAVADVGVQVQCYLGEPKTPWT 2036
Db 2090 AASEYVEVTQHGSFSVVTGLTSDNL---KVPCQVPAPEFFSWVDGVQIHRFAPTP----- 2141
Qy 2037 TSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALV 2096
Db 2142 ------GPFFRD---EVTFTVGLNSFVVGSQLPCDPEPDTEVLASMLTDPSHITAEAA-- 2190
Qy 2097 FKQELRRTNQLLEAISAGVDTTKLPAPSIEEV----------VVRKRQFRARTGSLTLP 2145
Db 2191 -ARRLARGSPPSQASSS---ASQLSAPSLKATCTTHKMAYDCDMVDANLFMG--GDVTRT 2244
Qy 2146 PPPRSVPGVSCPESL------QRSDPLE------GPSNLPP------SPPVLQLAMPM 2185
Db 2245 ESDSKVIVLDSLDSMTEVEDDREPSIPSEYLLKRRKFPPALPWAREYNPPVIE-TWKR 2303
Qy 2186 PLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKKE------ 2221
Db 2304 P----DYEPPTVLGCALPPTPQAP--VP--PPRRRRAKVLTQDNVEGVLREMANKVLSP 2354
Qy 2222 VSEWSDESWST-ATTASSYVTGP------PYPKIRGK----DSTQSAPA 2259
Db 2355 LQDHNDSGHSTGADTGDSVQQPSDETAVSDTGSLSMPPLEGEGDPLEFDPTGSAPP 2414
Qy 2260 KR------PTKKKLGKSEFSCSMSYTWTD-VISFKTASKVLSATRAITSGFLK 2305
Db 2415 SEGECEVIDSDSKSWSTVSDQEDSVICCSMSYSWTGALITPCGPEEEKLPINPLSNSLMR 2474
Qy 2306 QRSLVVVTEPRDAELRKQKVTINRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHT 2365
Db 2475 FHNKVYSTTSRSASLRAKKVTFDRVQVLDTHYDSVLQDVKRAASKVSARLLSVEEACALT 2534
Qy 2366 PSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEAGEIPSHYR-QTVIVPKEVFVKTPQ 2424
Db 2535 PPHSAKSRY-GFGAKEVRS-LSRRAVNHIRSVWEDLLEDQHTPIDTIVMAKNEVFCVDPT 2592
Qy 2425 KPTKKKPPRLLISYPHLEMRCVEKMYYGGVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMW-- 2481
Db 2593 KGGKKPARLIVTPDLGVRVCEKMALVDIAQKLPKAIMGPSYGFQYSPAERVDFLLKAWRS 2652
Qy 2482 SPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYD 2541
Db 2653 KKDPMGFSYDTRCFDSTVTERDIRTEESIYQACSLPQBARTVIHSLTERLYVGGPMTNSK 2712
```

```
Db  683 --------YSDLPALSTGLLHLHQNIVDVQYLYGLSPAVTRYIVKWEWVLLFLLLAD 732
Qy  593 SKYLIIAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLR 652
Db  733 AR--ICACLWMLI----------------------ILAQAEAALEKLI--- 757
Qy  653 CRLRYAALLGFVPMAAGLPLTFFVAAAAQPDYDWMVRLLVAGLVLWAGRNRGHRIALLV 712
Db  758 -------LHSASASANGPLWFFIFFTAA----WYLKGRVPVATYS--------VL 795
Qy  713 GPWPLVALLTLLHLVTPASAFDTEIG-GLTIPPVVALVVMSRFGFFAHLLPRCALVNS 770
Db  796 GLWSF--LLIVLALPQQAYALDIASQGELGLAIIAIISIFLT--PAYKILLSRSVWWLS 851
Qy  771 YLW------QRW-------ENWFW-NVTLRPERFFLV---LVCFPGATYDALVTFC 809
Db  852 YMLVLAEAQIQQWVPPLEARGGRDGLIWVAVILHPRLVFEVTKWLLALLGPAY------- 904
Qy  810 VCHVALLCLTSSAASPFGTDSRVRAHRM----LVRLGKCHAWYSHYVLKFPLLVFGE-N 863
Db  905 ----LLKASLLRVPYF-----VRAHALLRRMCTLVR----HLAGARYV-QMLLITMGRWT 949
Qy  864 GVFFYKHLHGDVLPNDFASKLPLQ------EP--FFPPEGKARVYRNEGRRL 907
Db  950 GTYIYDHLS---------PLSTWAAQGLRDLAVAVEPVVFSPMEKKVIVWGAE--TV 995
Qy  908 ACGDTVDGLPVVARLGDLVFAGLAMPPD----GWAITAPFTLQCLSERGTLSAMAVVMT 962
Db  996 ACGDIIHGLPVSARLGREVLLG---PADSYTSKGWKLLAPITAYTQQTRGLLGAIVSLT 1052
Qy  963 GIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHGSKGRRLAHPTGSIHPITVDAANDQ 1022
Db 1053 GRDKNEQAGQVQVLSSVTQSFGTSISGVLMWTVYHGAGNKTLAGPKGPVTQMYTSAEGDL 1112
Qy 1023 DIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAP 1082
Db 1113 VGWPSPPGTKSLDPCTCGAVDLYLVTRNADVIPVRRKDDRRGALLSPRPLSTLKGSSGGP 1172
Qy 1083 ILCSSGHVIGMFTA---ARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDYKPTVPNEYSV 1139
Db 1173 VLCSRGHAVGLFRAAVCARGVAKSIDFIPVESLNIATRTPSFSDNS---TPPAVPQSYQV 1229
Qy 1140 QILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCT 1199
Db 1230 GYLHAFTGSGKSTKVPAAYASQGYKVLVLNPSVAATLGFGAYMSKAHGINPHIRTGVRTV 1289
Qy 1200 NTGASLTYSTYGMYLT-GACSRN-YDVIICDECHATDATTVLGIKGVLTEAPSKNVRLVV 1257
Db 1290 TTGDSITYSTYGKFIADGGCSAGAYDVIICDECHSVDATTILGIGTVLDQAETAGVRLVV 1349
Qy 1258 LATATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELAN 1317
Db 1350 LATATPPGVTVTTPHSNIEEVALGHEGEIPFYGKAIPLAFIKGGRHLIFCHSRKKCDELAA 1409
Qy 1318 ELARKGITAVSYYRGCDISKIP-EGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVD 1376
Db 1410 ALRGMGVNAVAYYRGLDVSVIPTQGDVVVVATDALMGYTGDFDSVIDCNVAVTQIVDFS 1469
Qy 1377 LDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVVGDSCCTPSGMVPECNIVEARDAAK 1436
Db 1470 LDPTFTITTQTVPQDAVSRSQRRGRTGRGRLGIYRYVSSGERPSGMFDSVVLCECYDAGA 1529
Qy 1437 AWYGLSSTEAQTILDTYRTQPGLPAIGANLLDEWADLFS-MVNPEPSFVNTAKRTADNYVL 1495
Db 1530 AWYELTPAETTVRLRAYFNTPGLPVCQDHLEFWEAVFTGLTHIDAHFLSQTKQGGDNFAY 1589
Qy 1496 LTAAQLQLCHQYGVAAPNDAPRWQGARLGKKPCGVLMRLDGDAC------PGPEP-- 1545
Db 1590 LTAYQATVCAR---AKAPPPSWD----------VMWK------CLTRLKPTLTGPTPLL 1629
Qy 1546 ---------SEVTRY----QMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVR 1588
```

```
Qy   2542 GREIGYRRCRSSGVYTSSSNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGAD 2601
          |:. |||||||| .||:| |:.|||| .|| .:|..  .||:|| ||||| ||.:| ||
Db   2713 GQSCGYRRCRASGVFTTSMGNTMTCYIKALACKAAGIMDPIMLVCGDDLVVISESCQNE 2772

Qy   2602 ADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSNVTSGITKSGKPYYFLTRDPRI 2661
          .|: .:|||.: | |.|. ..||| |:.|||:||||||||.:.. .||| |||||||| :
Db   2773 EDERNLRVFTEAMTRYSAPPGDLPRPEYDLELITSCSSNVSVALDPRGRRRYFLTRDPTT 2832

Qy   2662 PLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGK 2721
          |: |.:|  .||| :||:||  |:.|| :||.|:. |:|| :.| |..| .:::.|:||.
Db   2833 PITRAAWETVRHSPVNSWLGNIIQYAPTIWVRMVIMTHFFSILLAQDTLNQNLNFEMYGA 2892

Qy   2722 NYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLAS 2781
          .|:|.: ||| || ..  .:.||..| :..|.|::.. |.:.||||||||.. |||||.|
Db   2893 VYSVNPLDLPAIIERLRHGLDAFSLHTYSPHELSRVAATLRKLGAPPLRAWKSRARAVRAS 2952

Qy   2782 AKRRGGAHAKLARFLL-WHATSR----PLPDLDKTSVARYTTFNYCDVYSPEGDVFITPQ 2836
          .:. ||.|.:  |.|| ||.||:    .|:|||.||:.|||:|||    :|.:|||| ||
Db   2953 LIAQGGRAATCGRYLFNWAVKTKLKLTPLPEASRLDLSGWFT-----VGAGGGDIFHSVS 3007

Qy   2837 RRLQKFLVKYLAVIVFALGLIAVGLAI 2863
          .| :||||.||| |.|.|||| ||| |
Db   3008 HARPRLL------LLCLLLLSVGVGI 3027
```

RESULT 10
Q7T7J1
```
ID   Q7T7J1      PRELIMINARY;     PRT;   3033 AA.
AC   Q7T7J1;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tanabe Y., Nagayama K., Enomoto N., Izumi N., Tazawa J., Sato C.,
RA   Watanabe M.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY232731; AAP55686.1; -.
HSSP; Q8JYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_Pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3033 AA;  330861 MW;  C6673E5964AEC019 CRC64;
```

```
Query Match             21.5%;  Score 3300.5;  DB 2;  Length 3033;
Best Local Similarity   29.8%;
Matches  982;  Conservative 444;  Mismatches 1163;  Indels 703;  Gaps 111;
                                        Pred. No. 2.6e-205.

Qy    8 TSPVPAPRTRKNKQTQASYPVSIK----------------------TSVERGQRAKRKVQRDARPR 51
        |:|:|:. :||:||.  :||||:|                      :|:|.|||:|:|.
Db    3 TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR 59

Qy   52 NYKIAGIHDGLQTLAQAALPAH------------GWRQDPRHKSRNL 87
        : :.. |:||||.| . .||.|            ||||||||:|||:
Db   60 GRRQPIPKDRRSTGKSWGKPGYPWPLYGNEGCGWAGWLLSPRGSRPHWGPTDPRHRSRNF 119

Qy   88 GILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLFV 144
        |:: | . .|:|.:  .||:| |.. .|.|.|.:|::||:||||||||:   .|. .:|:
Db  120 GKVIDTITCGFADLMGYIPVIGAPV-GGVARALAHGVRVLEDGVNYATGNLPGCSFSIFL 178

Qy  145 VCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVICADE------ 198
        : ||| | .|.|..  |: :.:. :.: |.|.:.:|  . ..:.|.||:.||..
Db  179 LALLSCVTVPVSAVEVRN-ISSSYYATNDCSNNSITWQLENAVLHLPGCVPCENDNGTLR 237

Qy  199 CWPANPYIS--HPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV-- 254
        ||..:.  :  ||| |||:| |||.|:|:|.||||||||||||||||||||||||:
Db  238 CWTQVTPNVAVKIRGALT---QHLRTHVDVIVVAATVCSALYVGDVCGAVMIASQALIVS 294

Qy  255 --RHWLIHIDLNETGTCYLEVPTGIDPGFLGPIGW--MAGKVEAVIFLTKLASQVPYAIA 310
          .||  |:|:|:||.|:.:  |.|..|.  |.|:  |.||:.|:.   |..:.|.
Db  295 PERH------NFTQECNCSIYQGRITGH--HMAWDMDMNLNWSPTITMIAYAARIPELVL 345

Qy  311 TMFSSVHYLAVGALIYYASRGKWYQLLLALMLY--IEA---TSGNPI 352
        |:|..||..::|.|.::.:...|.:: : .:.|  |||
Db  346 EVIFGGHWGVMFGLAYFSMQGAWSKVIVILLLVAGVDARHHTGLQVGKTLARVTSLFSI 405

Qy  353 ------------RVPTGCS---------IAEF------CSPLMIPC--------- 371
                    |||:||.         ||.|      ||.:: ||
Db  406 GPKQNIQLINTNGSWHINRTALNCNDSLQTGFIASLFVNNINSSGCPERMSSCRELDDF 465

Qy  372 --------------------PCHSYLSENV-SEVICYSPKWTRPITLEYNN 401
                            ||.|:|.|.| |||:||||.|.||:||||:|
Db  466 RIGWGTLEYETNVTNDEDMRPYCWHYPPKPCGIVPARTVCGPVYCFTPS---PIVVGTD 522

Qy  402 ------SISWYPYTIPGARGCMVKFKNNT-------WGC----------CRIR-- 431
              :||:|.:.|.|:.||  ||:|:|:       ||.          |||:
Db  523 KQGVPTYSW------GENETDVFLLNSTRPPRGSWFGCTWMNGTGFTKTCGAPPCRIRRD 576

Qy  432 -NVPSYCTMGTDAVVWNDTRNTYEACGVTPWLT------------TAWHNGSALKLAILQ---Y 478
         |.||||:. |:. .|:||.:.||:||| .|||             |.|||.|:|.|||:   |
Db  577 YNSTLDLLCPTDCFRKHPDTTYLKCGSGPWLTPKCLVEYPYRLWHYPCTVNFTIFKVRMY 636

Qy  479 PGSKEMFKPHNVMSGHLYF------------EGSDTPIVYFYDVNSTLLPPERWARLPGT 527
        .|. |:|:||||:|||| |            :.|||.||.||:|.||:  .||.|:||.|
Db  637 VGGVE----HRFSAACNFTRGDRCRLEDRDRGQQSPLLH------ST----TEWAVLPCS 682

Qy  528 PPVVRGSWLQVPQGFYSDVKDLIATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVLIL 587
        .|:|.:|.|:|.|.|:||||...  |||.:|:..|:|:||.:.|:||:|:..|::|.||:|
Db  683 ------FSDLPALSTGLLHLHQNIVDQVQYLYGLSPAITRYIVKWEWVLLF 727
```

```
Qy  1684 VSIIGR-----IHLNDRAVVAP--DKEVLYEAPDEMEECASKAALIEBGQRMAEMLKSKIQ 1737
Db  1636 ----TTTSPFTLETALE----KLNTFLGPHAATILAIIEYCCGLVTLPDNPPASCVFAF  1686

Qy  1738 GLLQQATRQAQDIQPAIQSSWPKLEQFWAKHMNFISGIQYLAGLSTLPGNPAVASMMAF   1797
Db  1687 IAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAGTALGTWTSVG---FVFD  1743

Qy  1798 SAALTSPLPTSTTILLNIMGGWLASQIAPPAGATGFVVSGLVGAAVG--SIGLGKILVD  1854
Db  1744 MLGGYAAASSTACLTFKCLMGGWPTMDQLAGLVYSAFNPAAGVGVGVLSACAMFALTTAGP 1803

Qy  1855 VLAGYGAGISGALVAFKIMSGEKPSVEDVVNLLPAILSPGALLVVGVICAAILRRHVGGGE 1914
Db  1804 D--HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVLSACIRWLHTPTED  1861

Qy  1915 GAVQWMNRLLIAFASRGNHVAPTHYVTESDASQRVTQVLSSLT---ITSLLRRLHTWITE  1970
Db  1862 DCGLI---AWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGM  1918

Qy  1971 DCPVPCSGSWLQDIDWDVVCSILTDFKNWLSS--KLLPKMPGLPFISCQKGYRGVWAGTGV 2028
Db  1919 LQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARP-DPTDWTSLVV  1977

Qy  2029 MTTRCPCGANISGHVRMGTMKI-TGPKTCLNLWQGTFPINCYTEBGPCVPKPPPNYKTAIW 2087
Db  1978 NYGVRDYCKYEKMGDHIFVTAVSSPNVCFTQVP---PTLRAAVAVDGVQVQCYLGEPKTP  2034

Qy  2088 RVAASEYVEVTQHGSFSYVTGLTSDNL---KVPCQVPAPEFFSWVDGVQIHRPAPIP---  2141
Db  2035 WTTSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSD--EAA   2091

Qy  2142 ------GPFFRD---EVTFTVGLNSFVVGSQLPCDPEPDTEVLASMLTDPSHITAEAA    2190
Db  2092 VSALVFKQELRRTNQLLEAISA-------------------------GVDTTKLPA--   2122

Qy  2191 ARRLARGSPPSQASSSASQLSAPSLKATCTTHKMAYDCDMVDANLFMGGDVTRIESDSKV 2250
Db  2123 ----------------PSI-EEVVVRKRQFRARTGSLTLPP--PRRSVPGVSCP--ESL 2160

Qy  2251 IVLDLSDSMTEVEDDREPSVPSEYLIRRRKF--------PPALPPWARPDYNPPVIETW  2301
Db  2161 QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK 2220

Qy  2302 KR------PGYEPP----------------------------TVLGCALPPTPQAP--VP--PPRR 2329
Db  2221 E-----------------------------------------VSEWSDESWST----ATTASSY 2239

Qy  2330 RRAKVLTQDNVEGVLREMADKVLSPLQDYNDSGHSTGADTGADSVQQPSDETAASEAGSL  2389
Db  2240 VTGPPYPKIRGKDSTQSAPAKR-----------PTKKKLGKSEFSCSMSVTWT        2281

Qy  2390 SSMPPLEGEGPDPDLEFEPARSAPPSEGECEIIDSDSKSWSTVSDQEDSVICCSMSYSWT  2449
Db  2282 D-VISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ 2340

Qy  2450 GALITPCGPEEEKLPINPLSNSLMRFHNKVYSTTSRSASLRAKKVTFDRVQVLDTHYDSV 2509
Db  2341 VRLAKEKASKVVGVWMDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA 2400

Qy  2510 LQDVKRAASKVSARLLSVEEACALTPPHSAKSRY-GFGAKEVRS-LSRRAVNHIRSVWED 2567
Db  2401 GEIPSHYR-QTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVERMYYGQVAPDVVKA 2459

Qy  2568 LLEDQHTPIDTTIMAKNEVFCVDPTKGGKKPARLIVPDLGVRVCEKMALYDIAQKLPKA  2627
Db  2460 VMGDAYGF-VDPRTRVKRLLSMM--SPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKL 2516

Qy  2628 VMGPSYGFQYSPAERVDFLLKAWGSKKDPMGFSYDTRCFDSTVTERDIRTEESIYQACSL 2687
Db  2517 SDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQ 2576
```

```
Qy   588 LGLCGSKYLILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFI   647
Db   728 LLLADAR--VCACLWMLI-----------------ILGQAEAALEKLII--           757

Qy   648 CCYLRCRLRYAALLGFVPMAAGLPLTFFVAAAAAQPDYDWWVRLLVAGLVLWAGRNRGHR   707
Db   758 -------------LHSASAASANGPLWFFIFFTAA-----WYLKGRVVPVATYS-       793

Qy   708 IALLVGPWPLVALLTLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGF---FAHLLPR   764
Db   794 ---VLGLWSF--LLLVLALPQQAYALDATDQGELGL--VMLVIISIFTLTPAYKILLSR  845

Qy   765 CALVNSYLW--------QRW--------ENWFW-NVTLRPERFFLV--LVCPFGGATYD    803
Db   846 SLWWLSYMLVLAEAQIQQWVPPLEARGGRDGIIWLAVILHPRLVFEITKWLLAILGPAY-  904

Qy   804 ALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRM---LVRLGKCHAWTYSHYVLKFFLL   858
Db   905 ---------------LLKMSLLRVPYF-----VRAHALLRVCTLVR----HLAGARYIQMLLIT  944

Qy   859 VFGENGVFFYKGHLHGDVLPNDFASKLPLQ----------EP--FFPPEGKARVYRN   902
Db   945 IGRWTGTYIYDHLS-------------PLSTWAAQGLRDLAVAVEPVIFSPMEKKVIVWGA  992

Qy   903 EGRRLACGDTVDGLPVVARLGDLIVFAGLAMPPD------GWAITAPFTLQCLSERGTLSAM  957
Db   993 E--TVACGDILHGLPVSARLGREVLLG---PADSYTSKGWKLLAPITAYTQQTRGLLGAI 1047

Qy   958 AVVMTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVD 1017
Db  1048 VVSLITGRDKNEQAGQVQVLSSVTQSFLGTSISGVLMTVYHGAGNKTLAGPKGPVTQMYTS 1107

Qy  1018 AANDQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKG 1077
Db  1108 ADGDLVGWPSSPGTKSLDPCTCGAVDLYLVTRNADVIPVRRKDDRRGALLSPRPLSTLKG 1167

Qy  1078 SSGAPILCSSGHVIGMFTA---ARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVP 1134
Db  1168 SSGGPVLCPRGHAVGLFRAAVCARGVAKSIDFIPVESIDIATRTPSFSDNS---TPPAVP 1224

Qy  1135 NEYSVQILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYF 1194
Db  1225 QSYQVGYLHAPTGSGKSTKVPAAYASQGYKVLVLNPSVAATLGFGAYMSKAHGINPNIRT 1284

Qy  1195 NGKCTNTGASLTYSTYGMYLT-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKN 1252
Db  1285 GVRTVVTGDSITYSTYGVGKFIADGGGCSAGAYDVIICDECHSVDATTILGIGTVLDQAETAG 1344

Qy  1253 VRLVVLATATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHC 1312
Db  1345 ARLVVLATATPPGTVTTPHSNIEEVALGHEGEIPFYGKAIPLAFIKGGRHLIFCHSKKKC 1404

Qy  1313 DELANELARKGITAVSYYRGCDISKIP-EGDCVVVATDALCTGYTGDFDSYDCSLMVEG  1371
Db  1405 DELAAALRGMGVNAVAYYRGLDVSVIPTQGDVVVVATDALMTGYTGDFDSVIDCNVAVTQ 1464

Qy  1372 TCHVDLDPTFMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEA  1431
Db  1465 IVDFSLDPTFTITTQTVPQDAVSRSQRRGRTGRGRLGIYRYVSSGERPSGMFDSVVLCEC 1524

Qy  1432 FDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTA 1490
Db  1525 YDAGAAWYELTPAETTVRLRAYFNTPGLPVCQDHLEFWEAVFTGLTHIDAHFLSQTKQGG 1584

Qy  1491 DNYVLLTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGDADAC------PG 1542
Db  1585 DNFAYLTAYQATVCAR---AKAPPPSWD--------VMWK------CLTRLKPTLTG  1624

Qy  1543 PEP---------SEVTRY--QMCFTEVNTSGTAALAVGVAMAYLAIDTFG  1583
Db  1625 PTPLLYRLGAVTNEVTLTHPVTKYIATCMQADLEIMTS-TWVLAGGVLAAVAAYCLATGC 1683

Qy  1584 ATCVRRCWSITSVPTGATVAPVVDEEBI------VEBCASFIPL-EAMVAAIDKLKSTI- 1635
```

```
Db   2688  PQEAKTVIHSLTERLYVCGPMTNSKGQSCGYRRCRASGVFTTSMGNTMTCYIKALAACKA  2747
                 |||||| | | |:||  ||||:|:|
Qy   2577  AGMKNPRFLICGGDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTS  2636
                 || : ||  ||:||||||                :|::| |    |: ||| :|
Db   2748  AGIMDPTMLVCGGDDLVVISESGQNEEDERNLRAFTEAMTRYSAPPGDIPRPEYDLELITS  2807

Qy   2637  CSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVL  2696
                 |||||:  :  : ||||| |||||:||  ||:|:  :||:|:|:||:|
Db   2808  CSSNVSVALDSRGRRRYFLTRDPTTPITRAAWETVRHSPVNSWLGNIIQYAPTIWVRMVI  2867

Qy   2697  AVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLLPSIIAGVHGIEAFSVVRYTNAEILRV  2756
                 |  | :    ||  :|| |  ||:||||| | | || :| |:|  || |:  || :|
Db   2868  MTHFFSILLAQDTLNQNLNFEMYGAVYSVNPLDLPAIIERIHGLDAFSLHTYSPHELSRV  2927

Qy   2757  SQSLTDTMPPLRAWRKKARAVLASAXRGGAHAKLARFLL-WHATSR----PLPDLDKT  2811
                 : ||||||:||||  |:| | : |||  ::|::  | |::   |:  |||  |:|
Db   2928  AATLRKIGAPPLRAWKSRARAVRASLIAQGGRAAICGRYLFNWAVKTKLKLTPLPEASRL  2987

Qy   2812  SVARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAI  2863
                 :::  ||   |:  :: ||:|:|| :|:
Db   2988  DLSGWFT----VGAGGGDIYHSVSHARPRLL-------LLCLLLLSVGVGI  3027


RESULT 11
Q9QAX1
ID   Q9QAX1      PRELIMINARY;  PRT;  3033 AA.
AC   Q9QAX1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=genotype 2; TISSUE=Serum;
RX   MEDLINE=20329282; PubMed=10872881; DOI=10.1023/A:1008182901274;
RA   Samokhvalov E.I., Hijikata M., Gylka R.I., Lvov D.K., Mishiro S.;
RT   "Full-genome nucleotide sequence of a hepatitis C virus variant
RT   (isolate name VAT96) representing a new subtype within the genotype 2
RT   (arbitrarily 2k).";
RL   Virus Genes 20:183-187(2000).
DR   EMBL; AB031663; BAA88057.1; -.
DR   HSSP; Q8JYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD_.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
```

```
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00271; Helicase_C; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
FT   CHAIN      734   1010       NS2 protein.
FT   CHAIN     1011   1661       NS3 protein.
FT   CHAIN     1662   1715       NS4A protein.
FT   CHAIN        1    191       core protein.
FT   CHAIN     1716   2017       NS4B protein.
FT   CHAIN     2018   2442       NS5A protein.
FT   CHAIN     2443   3033       NS5B protein.
FT   CHAIN      192    383       E1 protein.
FT   CHAIN      384    733       E2/NS1 protein.
SQ   SEQUENCE   3033 AA;  329860 MW;  524F9E0A46A7B50 CRC64;

Query Match             21.4%;  Score 3296.5;  DB 2;  Length 3033;
Best Local Similarity   30.4%;  Pred. No. 4.8e-205;
Matches  985;  Conservative 466;  Mismatches 1191;  Indels 597;  Gaps 115;
```

```
Qy     8  TSPVPAPRTRKN--KQTQASYP-----------------VSIKTSVERGQRAKRKVQR   46
             |:|:| ||   ||        :|            : :||   :
Db     3  TNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRGRR   62

Qy    47  DARPRNYKIAGI--------------------HDGLQTLAQAALP--AHGWGRQDPRHKSRNLGIL   90
             |:|:                         | | || ||||   :|: :|:||||||||:
Db    63  QPIPKDRRSAGKSWGRPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPTDPRHRSRNLGKV  122

Qy    91  LDYPLGWIGVDTVTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG--WFGVHLFVVCL  147
             :| || ||||:|||||||:|||||||||||||| ||||||||||    | |:|:|||:
Db   123  IDTLTCGFADLMGYIPVVGAPV-GGVARALAHGVRVLEDGINYATGNLPGCSFSIFLLAL  181

Qy   148  LS-LACPCSGARVTDPDTNTT----ILTNCCQRNQVIYCSPSTCLHEPGCVIC----ADE  198
             |: |:|  |||  | :|:|     ||  | | |||| : | :|:|:|
Db   182  LSCMSVPVSAVEV---KNTSQIYMATNDCSNNSITWQLEGAVLHVPGCVPGCESTGNISR  237

Qy   199  CWPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELGACVLVGDWLV--R  255
             |:|   | :| | ||| : ||:| |:|:||||| ||||| :||||  | :|
Db   238  CWIPVTPNVAVRRRGALTKG-LRTHIDLIVVSATFCSALYIGDVCGAIMIAAQATIISPQ  296

Qy   256  HWLIHIDLNETGTCYLEVPTGIDPGFL--GFIGW--MAGKVEAVIFLTKLASQVPYAIAT  311
             |  |:|:|:| ||| |||||||||:|  ||:|   ||||| |:| || |||||||:|
Db   297  HHTFVQDCN-------CSIYPGHVTGHRMAWDMMMWNWSPATTMIMAYFMRVPEVVLD  346

Qy   312  MFSSVHYLAVGALIYYASRGKWYQLLLALML-----------------------  342
             | :|| | |:|  | ||  :|:
Db   347  IITGAHWGVMFGLAYFSMQGAWAKVVVLLLTAGVDAQTHTISGHAARTTHGLVSLFTPG  406

Qy   343  ----------------------YIEA-----------TSGNPIRVPTGCSIAEF-  363
                                   ::|            ||||| :| ||| | :|
Db   407  SQQNIQLVNTNGSWHINRTALNCNDSLKTGFIAALFYSHKFNSSGCPQRMSSCRSIEEFR  466

Qy   364  ------------------C--SPLMIPCPCHSYLSENV-SEVICYSPKWTRPITLEY  399
                               |     |:  ||| || :|| |  ||
Db   467  IGWGNLEYEENVTNDDNMRPYCWHYP---PRPCGIVPAQTVCGPVYCFTPS---PVVVGT  520

Qy   400  NNSISWVPYTIPGARGCMVKFKNNT--------------------CRIR---NV  433
             :  |:|  |||    | |:| |||                     |:|:   ||
Db   521  TDRRGVPTYTW-GENDTDVFLLNSTRPPRGAWFGCTWMNSTGFTKTCGAPPCRIRPDFNS  579
```

```
Qy              .........                                                                           QM 1553
1542 Db  RLRAYFNTPGLPVCQDHLEBFWEAVFTGLTHIDAHFLSQTKQAGENFPYLVAYQATVCARA 1601

1508 Qy  GYAAPNDAPRWQG-ARLGKK----PCGVLWRLDGADACPGPEPSE------VTRY--QM 1553
1602 Db  KAPPPSWDVMWKCLIRLKPTLTGPTPLLYRL------GPVTNETTLTHPVTKYIATCMQ 1654

1554 Qy  CPTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI-- 1611
1655 Db  ADLEIMTS-TWVLAGGVLAAIAAYCLATGCVVCIGR---VNINQKTIVAP-DKEVLYE 1707

1612 Qy  ----VEECAS-FIPLEAMVAAIDKLKSTI-----TTTSPFTLETALE----KLNTFLGP 1656
1708 Db  APFDEMEECASRALLLEBGQRIAEMLKSKIQGLLQQATKQAQDIQPAVQATWPKLEQFWAK 1767

1657 Qy  HAATILAIIEYCCGLVTLPDDNPPASCVFAPIAGITTPLPHKIKMFLSLFGGAIASKLTDA 1716
1768 Db  HMWNFISGIQYLAGLSTLPGNPAVAAMMAFSAALTSPLPTSTTILLNIMGGWLASQIAPA 1827

1717 Qy  RGALAFMMAGAAGTALGTWTSVG---FVFDMLGGYAAASSTACLTFKCLMGEWPTMDQLA 1773
1828 Db  AGATGFVVSGLVGAAVG--SIGLGKILVDVLAGYGAGISGALVAFKIMSGBKPSVEDVV 1884

1774 Qy  GLVYSAFNPAAGVVGVLSACAMFALTTAGPD-HWPNRLLTMLARSNTVCNEYFIATRDI 1831
1885 Db  NLLPGIISPGALVVGVICAAILRRHVGQGEGAVQWMNRLIAFASRGNHVAPTHYVAESDA 1944

1832 Qy  RRKILGILEASTPWSVISACIRWLHTPTEDDCGLI---AWGLEIWQYVCNFFVICFNVLK 1888
1945 Db  SQRVTQLLGSLT----ITSLLRRLHTWITEDCPVPCAGSWLRDIWDWACTILTDFKNWL- 1999

1889 Qy  AGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCS 1948
2000 Db  -STKLLPKMPGLPFISCQRGHKGAWTGTGIMTTRCPCGAVVSGNVRHGSMRI-TGPKTCM 2057

1949 Qy  NYWRGAVPVNARLCGSARPDPT-DWTSLVVNYGVRDYCYEKMGDHIFVTAVSSPNVCFT 2007
2058 Db  NTWQGTFPINCYTEGQCAPQPTHNYKTAIWKVAAAEYAEVTRHGSYAYVTGLTNDNL--- 2114

2008 Qy  QVPPTLRAAVA---VDGVQVQCYLGEPKTPWTTSACCY--GPDGKGKTVKLPPRVDGHTP 2062
2115 Db  KVPCQLPAPEFFSWVDGVQIHRFAPTPK-PFIRDEVSPTVGLNSFVVGSQLPCEPEPDTE 2173

2063 Qy  GVRMQLNLRDALETNDCNST------NVTPSDEAAVSALVFKQELRRT----NQ 2106
2174 Db  VLASML-------TDPSHITAEAAAARRLARGSPPSEASSSASQLSAPSLRATCTAHAKNY 2226

2107 Qy  LLEAISA----GVDTTKLPA--------PSIE------EVVVRKRQFRARTG 2140
2227 Db  AVEMVDANFFMGSDVTRIESETKVLILDSLDPSVEBEDEREPSVPSEYLLPKKKKF---- 2281

2141 Qy  SLTLPPPPRSVPGVSCP--------ESLQRSD--PLEGPSNLPP----SPPVLQLAMP 2184
2282 Db  ------PQALPWARPDYNPPVVETWKRPDYDPPTVSGCALPPRVTAPTPPPRRRRALV 2334

2185 Qy  MPLLGAGECNPFTAI--------GC-AMTETGGGPDDLPSYPPKKEVSEWSDESWSTATT 2235
2335 Db  LSQSNVGEALQALAIKSFGQLPPSCDSGSRSTGMDTTDATDQPALKE--------STDSE 2385

2236 Qy  ASSYVTGPPYPKIRGKDSTQSA------PAKR--------PTKKKLGKSEFSCSMS 2277
2386 Db  AGSDSSMPPLEGEPGDPDLESSGSVEYHPSSQEGEAAPDLDSGSWSTCSEBEGGSEVCCSMS 2445

2278 Qy  YTWTD-VISPKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLPPPS 2336
2446 Db  YSWTGALITPCGPEEEKLPINPLSNSLLRYHNKVYSTTSRSASQRAKKVTFDRVQLLDSH 2505

2337 Qy  YHKQVRLAKEKASKVVGVWMDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQK 2396
2506 Db  YDQVLKDIKLAASKVSANLLSIEEACALITPPHSARSKY-GFGAKEVRS-LSRKAVDHIKS 2563

2397 Qy  ----CVEBAGEIPSHYRQTVIVPKEEVFVKTPQKPRLISYPHLEMRCVEKMYYGQV 2452
```

```
434 Qy  PSYCTMGTDAVWNDTRNTYEACGVTPWLT------TAWHNGSALKLA---ILQYPGS 481
580 Db  SEDLLCPTDCFRKHSEATYTRCGAGPWLTPKCLFHYPYRLWHYPCTINFFTHKIRMFIGG 639

482 Qy  KEMPKPHNWMSGHLYFBGSDTPIVYFYDPVNSTLLP---PERWARLPGTPPVVRGSWLQ 537
640 Db  VE----HRLEAACNFTRGDRCNL--EDRDRSQLSPLLHSTTEWAILPCT------- 682

538 Qy  VPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLI 597
683 Db  ----FSDMPALSTGLLHLHQNIVDVQYLYGLSPAITKYIVKWEWVVLLFLLLADAR--V 735

598 Qy  LAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRY 657
736 Db  CACL-WMLLLLGQA----------EAALEKLVILHAASAASSHGMLCP------- 772

658 Qy  AALLGFVVMAAGLPLTFVAAAAAQPDYDWWVRLLVAGLVLVWAGRNRGHRIALLVGPWPL 717
773 Db  ----------IFFFIAA--------WYIKGRVTPLVTYS--------YLGMWSF 800

718 Qy  VALLTLHLVTPASAFDTE----------IIGGLTIPPVVALVVMSRFGFPAHLLPRC- 765
801 Db  --SLLLLALPQQAYALDTTEQCQIGLVLLVVISVFTLSPAYKILLCRSLWWLSYLLVRAE 858

766 Qy  ALVNSYL-WQR---WENWFWNVTLRPERFFLVLVCFPGATYDALVTFCVCHVA----LL 816
859 Db  ALIQDWVPPWQARGGRDGIIWAAT------IFC-PGVLFD--ITNWLLAILGPGYLL 906

817 Qy  CLTSSAASFFGTDSRVRAHRML-----VRLGKCHAWYSHYLKFFLLVFGE-NGVFFYKH 870
907 Db  RSVLTSTPYF-----VRAQALLRICAAVR----HLSGGKYV-QMMLLTLGKWTGTYIYDH 956

871 Qy  LHGDVLP----------NDFASKLPLQEP--FPPFEGKARVYRNEGRRLACGDTVDGLPVV 919
957 Db  LS----PMSGWAAASGLRDLAVAV--EPIVFSPMEKKVIVWGAE--TAACGDILHGLPVS 1007

920 Qy  ARLGDLVFAGLAMPPD------GWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIF 974
1008 Db  ARLGQEVLLG---PADEYTSKGWKLLAPITAYAQQTRGLLGTIVVSMTGRDKTEQAGEIQ 1064

975 Qy  RLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDIYQPPCGAGSL 1034
1065 Db  VLSTVTQSFLGTTISGILWTVFHGAGNKTLAGSRGPVTQMYSSAEGDLVGWPSPPGTRSL 1124

1035 Qy  TRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMF 1094
1125 Db  DPCTCGAVDLYLVTRNADVIPARRQGDRRGALLSPRPLSSLKGSSSGGPVLCPRGHAVGIF 1184

1095 Qy  TAA---RNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILIAPTGSGKS 1151
1185 Db  RAAICTRGAAKSIDFIPIESLDVIIRSPNFTDNS---SPPAVPQTYQVGYLHAPTGSGKS 1241

1152 Qy  TKLPLSYMQBKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYG 1211
1242 Db  TKVPASYAAQGYKVLVLNPSVAATLGFGAYMSKAHGINPNIRTGVRTVTTGESITYSTYG 1301

1212 Qy  MYLT-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPT 1269
1302 Db  KFPLADGGCSGGAYDVIICDECHSVDATTILGIGTVLDQAETAGARLTVLATATPGSVTT 1361

1270 Qy  PHANITBIQLTDEGTIPPHGKKIKEENLKKGRHLIFEATKKHCDBLANELARKGITAVSY 1329
1362 Db  PHPNIEVALGHEGEIPFYGKAIPLSQIKGGRHLIFCHSKKKCDELAAALRGMGLNAVAY 1421

1330 Qy  YRGCDISKIP-EGDCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVC 1388
1422 Db  YRGLDVSVIPTGGDVVVVATDALMTGFTGDFDSVVDCNVAVTQTVDFSLDPTFTVTTQTV 1481

1389 Qy  GVSAIVKGQRGRGTRGRGAGIYYVVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQT 1448
1482 Db  PQDAVSRSQRRGRTGRGRGLGIYRYVSSGERASGMPDSVVLCECYDAGAAWYELTPAETTV 1541

1449 Qy  ILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLCHQY 1507
```

```
Db 2564 VWKDLLEDQQTPI---PTTIMAKNEVFCIDPTKGGKKAARLIVFPDLGVRVCEKMALYDI 2620
Qy 2453 APDVVKAVMGDAYGF-VDPRTRVKRLLSMW--SPDAVGATCDTVCFDSTITPEDIMVETD 2509
Db 2621 TQKLPQAVMGASYGFQYSPAQRVDFLLRAWKEKKDPMGFSYDTRCFDSTVTERDIRTEES 2680
Qy 2510 IYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVVTSSSNSLTCWLK  2569
Db 2681 IYLACSLPEEARVAIHSLTERLYVGGPMMNSKGQSCGYRRCRASGVLTTSMGNTIITCYVK 2740
Qy 2570 VNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKY 2629
Db 2741 ALAACKAAGIVAPTMLVCGDDLVVISESQGAEEDERNLRVFTEAMTRYSAPPGDPPKPEY 2800
Qy 2630 SLEELTSCSSNVTSGITKSGKPYYFLIRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPC 2689
Db 2801 DLELITSCSSNVSVALDQHGRRMYYLTRDPSTPLARAAWETARHSPVNSWLGNIIQYAPT 2860
Qy 2690 LWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYT 2749
Db 2861 IWVRMVLMTHFFSVLMAQETLDQDLNFEMYGAVYSVNPLDLPAIIERLHGLEAFSLHGYS 2920
Qy 2750 NABILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WHATSR----P 2804
Db 2921 PTELTRVAAALRKLGAPPLRAWKSRARAVRASLISQGGRAATCGFYLFNWAVRTKRKLTP 2980
Qy 2805 LPDLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAI  2863
Db 2981 LPAARRLDLSGWFT-----VGAGGGDIYHSVSRARPRFL------LLCLLLSVGVGI    3027
```

RESULT 12

```
Q7T7J2
ID   Q7T7J2        PRELIMINARY;     PRT;   3033 AA.
AC   Q7T7J2;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tanabe Y., Nagayama K., Enomoto N., Izumi N., Tazawa J., Sato C.,
RA   Watanabe M.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY233730; AAP55685.1; -.
DR   HSSP; Q8JYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002521; HCV_capsid.
DR   InterPro; IPR002522; HCV_core.
DR   InterPro; IPR002531; HCV_env.
DR   InterPro; IPR000745; HCV_NS1.
DR   InterPro; IPR001490; HCV_NS4a.
DR   InterPro; IPR002868; HCV_NS4b.
DR   InterPro; IPR002519; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS
DR   InterPro; IPR007094; RNA_pol_PSvir
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; Viral RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3033 AA;  330641 MW;  593E6BD2358AF44E CRC64;
```

```
Query Match             21.4%;
Best Local Similarity   30.1%;
Matches  984;  Conservative  448;  Mismatches 1176;  Indels  665;  Gaps 111;
Score 3293.5;  DB 2;  Length 3033;
Pred. No. 7.5e-205;

Qy    8 TSPVPAPRTRKNKQTQASYPVSIK------------------------TSVERGQRAKRKVQRDARPR 51
         :     ||:||:|||
Db    3 TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR 59

Qy   52 NYKIAGIHDGLQTLAQAALPAH----------------------GWGRQDPRHKSRNL 87
         :    :                                     ||:||||
Db   60 GRRQPIPKDRRSTGKSWGKPGYPWPLYGNEGCCWAGWLLSPRGSRPHWGPTDPRHRSRNL 119

Qy   88 GILLDYPLGWIGDVTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLFV 144
         |       : |:|:||  ::|:||  |: |    :|:||||||:|    WFGVHLFV
Db  120 GKVIDTITCGFADLMGYIPVIGAPV-GGVARALAHGVRVLEDGVNYATGNLPGCSFSIFL 178

Qy  145 VCLLS-LACPCSGARVTDPDTNTTILNCCQRNQVIYCSPSTCLHEPGCVICADE------ 198
         :   :  |||:| | : :   ||   :|
Db  179 LALLSCVTPVSAVEVRN-ISSSYYATNDCSNNSITWQLENAVLHLPGCVPCENDNGTLR 237

Qy  199 CWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDMLV---- 254
         :    ::   :    ::    : : :: |||:|: :   | ::|  ::|
Db  238 CWTQVTPNVAVKHRGALTQN-LRTHVDVIVVAATVCSALYGVDVCGAVMIASQALIVSPA 296

Qy  255 RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGW--MAGKVEAVIFLTKLASQVPYAIATM 312
         |       ::  ||  :||     || | |  :
Db  297 RH------NFTQECNCSIYQGRITGH--HMAWNDMLNWSPTITMILAYAARIPELVLEV 347

Qy  313 FSSVHYLAVGALIYYASRGKWYQLLLALMLY-IEA----TSGNPIRVPTGCSIAEFCSPL 367
         |:      :|   | ::|  ||:|||:||
Db  348 IFGGHWGVMFGLAYFSMQGAWAKVIVILLLVAGVDARHHTTG----LQAGKTLARVTSLF 403

Qy  368 MIPCP----------------CHSYLS------------------------ENVSEVIC 386
         :   :                :|   :                         ::::
Db  404 SIGAKQNIQLINTNGSWHINRTALNCNDSLQTGFIASLFYVNNINSSGCPERMSSCRELD 463

Qy  387 -YSPKWTRPITLEYNNSIS-----------W-YP----------------YTIPGARGCM 417
         :   ::        :          :||      :|               ::||  ||
Db  464 DFRIGWG---TLEYETNVTNDEDMRPYCWHYPPKPCGIVPARTVCGPVVCFTPSPIVVGT 520

Qy  418 -----------------------------------CRIR---NVP 434
         |||:     ||
Db  521 TDKQGVPTYSWGENETDVFLLNSTRPPRGSWFGCTWMNGTGFTKTCGAPPCRIRRDYNST 580

Qy  435 SYCTMGTDAVWNDTRNTVEACGVTPWLT--------------TAWHNGSALKLAILQ---YPGSK 482
         |   : ||   || |  : |||  |                : || : :: |:|
Db  581 LDLLCPTDCFRKHPDTTYLKCGSGPWLTPKCLVEYPYRLWHYPCTVNFTIFKVRMYGAV 640

Qy  483 EMFKPHNWMSGHLYF-----------EGSDTPIVYFYDPVNSTLLPPERWARLPGTPVV 531
         :    : :|  | |:            |  |||  :|:|  : :  |     | | |
Db  641 E----HRFSAACNFTRGDRCRLEDRDRQQSPLLH------ST----TEWAVLPCS---- 682
```

```
Qy   532  RGSWLQVPQGFYSDVKDLATGLIITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLC  591
          :| ||     : :||  | |: :         : :       |  |||   |::: |::
Db   683  -------FSDLPALSTGLLHLHQNIVDVQYLYGLSPAITRYIVKWEWVLLFLLLA       731

Qy   592  GSKYILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIPFICCYL    651
          ::                               :|: ||
Db   732  DAR---VCACLWMLI----------------------ILGQABAALEKLII--------- 757

Qy   652  RCRLRYAALLGFVPMAAGLPLTFVAAAAAQPDYDWWVRLLVAGLVLVLWAGRNRGHRIALL  711
                     ::  |||  |   || ||  :     : ||    :
Db   758  -------LHSASAASANGPLWFFIFFAAA-----WYLKGRVPVPVATYS-----:-:    794

Qy   712  VGPWPIVALLTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGF---FAHLLPRCALV   768
          :|| |:|  | :|::|     || : || |   ||    :        :||||   | :
Db   795  LGLWSF-LLLVLALPQQAYALDATDQGELGL---VVLVIISIFFLTPAYKILLSRSLWW    849

Qy   769  NSYLW-----QRW--------ENWFW-NVTLRPERFFLV---LVCFPGATYDALVT      807
           :|:       ||        | : |   : |  :|:        | ||| ||
Db   850  LSYMLVLAEAAQIQQWVPPLEARGGRDGIIWLAVILHPRLVFEITKWLLAILGPAY----- 904

Qy   808  FCVCHVALLCLTSSAASFFGTDSRVRAHRM------LVRLGKCHAWYSHYVLKFLLVFGE   862
           |    |  |     :::      ||  ||       |  |:       |::  |:
Db   905  -----LLKMSLLRVPYF-----VRAHALLRVCTLVR----HLAGARYIQMLLITIGRW    948

Qy   863  NGVFFYKHLHGDVLPNDFASKLPLQ------EP--FFPPEGKARVYRNEGRR          906
                                             |    |||| |||    ||
Db   949  TGTYIYDHLS-------------PLSTWAAQGLRDLAVAVEPVIFSPMEKKVIVWGAE--T  994

Qy   907  LACGDTVDGLPVVARLGDLVFAGLAMPPD------GWAITAPFTLQCLSERGTLSAMAVVM  961
          :|||| : ||||::  |  |  |  |:           ||::|   | |   | | ::  :
Db   995  VACGDILHGLPVVSARIGREVLLG---PADSYTSKGWKLLAPITAYTQQTRGLLGAIVVSL 1051

Qy   962  TGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAND 1021
          ||  :        | ||:| : | | | :||  |  |: | :  | ||:      :||| |
Db  1052  TGRDKNEQAGBEVQVLSSVTQSFLGTSISGVLWTVYHGAGNKTLAGPKGPVTQMYTSADGD 1111

Qy  1022  QDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVVCGALPMAVAKGSSGA 1081
          :|||| | |: ||  | || :| |||||   |:|:   || |||: ||:   :  ||| ||
Db  1112  LVGWFSPPGTKSLDPCTCGAVDLYLVTRNADVIPVRRKDDRRGALLSPRPLSTLKGSSSGG 1171

Qy  1082  PILCSSGHVIGMFTA--ARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYS  1138
          |:||  || :|:| |   |    ::: :  : |   | ||| || |:| |   || :|:|
Db  1172  PVLCPRGHAVGLFRAAVCARGVAKSIDFIPVESIDIATRTPSFSDNS---TPPAVPQSYQ 1228

Qy  1139  VQILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKC 1198
          | :|:|||||||||  | : : ::  :|||||||| | |   || : ||:||| : |:|
Db  1229  VGYLHAPTGSGKSTKVPAAYASQGYKVLVLNPSVAATLGFGAYMSKAHGINPNIRTGVRT 1288

Qy  1199  TNTGASLTYSTYGMYLT-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLV 1256
           :|| |:: |||| ::| |:|: : |||||||||| |||||:|||| ||  | |:||||
Db  1289  VTTGDSIITYSTYGKFIADGGCSAGAYDVIICDECHSVDATTILGIGTVLDQAETAGARLV 1348

Qy  1257  VLATATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKGHCDELA 1316
          |||||||||   ||| ||||:::    ||||  | |  :|  ||||||| | | |||||
Db  1349  VLATATPPGTVTTPHSNIEEVALGHEGEIPFYGKAIPLAFIKGGRHLIFCHSKKKCDELA 1408

Qy  1317  NELARKGITAVSYYRGCAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAA 1375
            | :  :  | :|||      : ||||| |  :|  ||        || |  : : ||
Db  1409  AALRGMCVNAVAYYRGLDVSVIPTQGDVVVVATADLMTGYTGDFDSVIDCNVAVTQIVDF 1468

Qy  1376  DLDPTFTMGVRVCGVSAIVKGVRRGRTGRGRLGIYRYVSSGERPSGMFDSVVLCECYDAG 1435
          |||||| : | : :  || | : || |  || || || : :   | |   ::   | || |
Db  1469  SLDPTFTITTQTVPQDAVSRSQRRGRTGRGRLGIYRYVSSGERPSGMFDSVVLCECYDAG 1528

Qy  1436  KAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYV 1494
           |||     | |::   | | |||| :  : : ||:       |||||||  | :  | :
Db  1529  AAWYELTPAETTVRLRAYFNTPGLPVCQDHLEFWEAVFTGLTHIDAHFLSQTKQGGDNFA 1588

Qy  1495  LLTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADAC-------PGPEP- 1545
           |||:| ::  | ||  | :  : : | |  :|| ||               ||||
Db  1589  YLTAYQATVCAR-----AKAPPPSWD---------VMWK------CLTRLKPTLTGPTPL 1628
```

```
Qy  1546  -------SEVTRY----QMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCV 1587
                 || |         ::|| |  :    : :: ||  | ||| :  | || ::
Db  1629  LYRLGAVTNEVTLTHPVTKYIATCMQADLEIMTS-SWVLAGGVLAAVAAYCLATGCVSII 1687

Qy  1588  RRCWSITSVPTGATVAPVVDEEI----VEECASFIPL-EAMVAAIDKLKSTI------ 1635
                          | :     |      ::|| |  |   :: |  :: ||
Db  1688  GR----IHLNDRAVVAP--DKEVLYEAFDEMEECASKAALIEEGORMAEMLKSKIQGLLQ 1741

Qy  1636  -TTTSPFTLETALE----KLNTFLGPHAATILAIIEYCCGLVTLPDNPFASCVFAFIAGI 1690
           ::: ||  | : |     ||: ::   || | :|:   |: ||| |||   ::  :: |:
Db  1742  QATRQAQDIQPAIQSSWPKLEQFWAKHMWNFISGIQYLAGLSTLPGNPAVASMMAFSAAL 1801

Qy  1691  TTPLPHKIKMFLSLFCGAIASKLTDARGALAFMMAGAAGTALGTWTSVG---FVFDMLGG 1747
           |:|||     ::| ||  |||:   | || || ||| ||  || |:| |     | ||
Db  1802  TSPLPTSTTILLNIMGGWLASQIAPPAGATGVVSGLVGAAVG---SIGLGKLVDVLAG 1858

Qy  1748  YAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTTAGPD--H 1805
          | | || | : |  :|| :|  : : ||||:| ||| ||||| ||| :| |  || |
Db  1859  YGAGISGALVAFKIMSGEKPSVEDVVNLLPAILSPGALVVGVICAALLRRHVGQGEGAVQ 1918

Qy  1806  WPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTEDDCGL 1865
          | :::    :||:  |   :   :  | :  : |  |::||| ||  |: | |:|:| ||
Db  1919  WMNRLIAFASRGNHVAPTHYVTESDASQRVTQVLSSLT----ITSLLRRLHTWITEDCPV 1974

Qy  1866  I---AWGLEIWQYVCNFFVICENVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQAR 1922
              :| |: | :||  : | : :|::    :|:|||||| ||||||:|  | |:|  |
Db  1975  PCSGSWLQDIWDWVCSILTDFKNWLSS--KLLPKMPGLPFISCQKGYGRVWAGTGVMTTR 2032

Qy  1923  CPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARP-DPTDWTSLVVNYGV 1981
          ||||| :|  :  | :| : || ||:| ||| ||  : ||:  |   ||    :   : |
Db  2033  CPCGANISGHVRMGTMKI-TGPKTCLNLWQGTFPINCYTEGPCVPKPPNYKTAIWRVAA 2091

Qy  1982  RDYCKYEKMGDHIFVTAVSSPNVCFTQVP---PTLRAAVADVGVQVQCYLGEPKTPWTTS 2038
           :| :|: : |  :|:   ||:| :     ::| :    || ||:|   :
Db  2092  SEYVEVTQHGSPSYVTGLTSDNL---KVPCQVPAPEFFSWVDGVQIHRFAPIP------ 2141

Qy  2039  ACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSD---EAAVSAL 2095
             ||| || || | |||:||  |:   :|   : | || | |:  : |:     ||: ||
Db  2142  ----GPFFRD---EVTFTVGLNSFVVGSQLPCDPEPDTEVLASMLTDPSHITAEAAARRL 2194

Qy  2096  VFKQELRRTNQLLEAISA----------GVDTTKLPA------ 2122
                   ::   : :|:          |:||  :
Db  2195  ARGSPPSQASSSASQLSAPSLKATCTTHKMAYDCDMVDANLFMGGDVTRIESDSKVIVLD 2254

Qy  2123  ------PSI-EEVVRKRQFRARTGSLTLPP--PPRSVPGVSCP-ESLQRSD 2164
                |:: : :   |  :  :|| |||   ||||   | |  ::|:
Db  2255  SLDSMTEVEDDREPSVPSEYLIRRKP-------PPALPPWARPDYNPPVIETWKR-- 2303

Qy  2165  PLEGPSNLPPS-------PPVLQLAMPML----------LGAGECNPFTAIG 2200
           |  |  ||||        || | |  || :          |||||| |
Db  2304  ----PGYDPPTVLICALPPTPQAPVPPPRRRRAKVLTQDNVEGVLREMADKVLSPLQDYN 2359

Qy  2201  CAMTETGG--GPDDLPSYPHGKKVSEWSDESWSTATTASSYVTGPPYKIRGKDSTQSAP 2258
            : |: |   || ||||  : : :: |||:|| ||| ||| ||| :    | |:||
Db  2360  DSGHSTGADTGAD----------SVQQPSDE--TAASEAGSLSSMPPLEGEPGDPDLEFEP 2408

Qy  2259  AKR-----------PTKKKLGKSEFSCSMSYTWTD-VISFKTASKVLSATRAI 2299
          |:            |||| ||| | | |||||| | ||| : || | | | |
Db  2409  ARSAPPSBEGECEIIDSDSKSWSTVSDQEDSVICCSMSYSWTGALIITPCGPEEEKLPINPL 2468

Qy  2300  TSGFLKQRSLVVTEPRDAELRKQKVTINRQLFPPSYHKQVRLAKBEKASKVVGVMWDYD 2359
          :|  |:: :  | | ||| | |  | :: : | |:| : | || :|||| | | : |
Db  2469  SNSLMRFHNKVYSTTSRSASLRAKKVTFDRVQVLDTHYDSVGTPSGMFDSVVSARLLSVE 2528

Qy  2360  EVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEAGEIPSHYR-QTVIVPKEEV 2418
          | |:||| |||||  ||| | :| |  |: :   || :: | :||:| | ||: : ::||
Db  2529  EACALTPPHSAKSRY-GFGAKEVRS-LSRRAVNHIRSVWEDLLEDQHTPIDTTIMAKNEV 2586

Qy  2419  FVKTPQKPTKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRL 2477
          | : | | || |||| ||:| :||||||  | ||| ||  :|| :|||   | : | ||
Db  2587  FCVDPTKGGKKPARLIVYPDLGVRVCEKMALYDIAQKLPKAVMGPSYGFQYSPAERVDFL 2646

Qy  2478  LSMW--SPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGG 2535
```

```
Db   2647 PKAWGSKKDPMGFSYDTRCFDSTVTERDIRTEESIYQACSLPQEAKTVIHSLTERLYVGG 2706
                :|: .|| |||||| ||     :  ||::   ||:  :|||||
Qy   2536 PMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIW 2595
           ||   ||    ||||| |||  :   :|:  | |:     :    |:  || |||:||||
Db   2707 PMTNSKGQSCGYRRCRASGVFTTSMGNTMTCYIKALAACKAAGIMDPTMLVCGDDLVVIS 2766

Qy   2596 KSAGADADKQAMRVFAGWMKVMGAPQDCVPQPKYSLEELTSCSSNVTSGITKSGKPYYFL 2655
           :||  | :: | |     ||  | ||: || |  |:  |:|||||::|  |:|||  |||
Db   2767 ESQGNEEDERNLRAFTEAMTRYSAPPGDIPRPEYDLELITSCSSNVSVALDSRGRRRYFL 2826

Qy   2656 TRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVT 2715
           |||| |  ||    |||    |  |||  :||||||||||:|| ||::| |||:|   :|
Db   2827 TRDPTTPITRAAWETVRHSPVNSWLGNIIQYAPTIWVRMVIMTHFFSILLAQDTLNQNLN 2886

Qy   2716 FDWYGKNYTVPVEDLPSIIAGVHGIEAPSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKA 2775
           |:  | |   |:     |||| | ||:    :|   ||||| :|| ||: |||:|||||:
Db   2887 FEMYGAVYSVNPLDLPAIIERIHGLDAFSLHTYSPHELSRVAATLRKLGAPPLRAWKSRA 2946

Qy   2776 RAVLASAXRRGGAHAKLARFLI-WHATSR-----PLPDLDKTSVARYTTFNYCDVYSPEGD 2830
           ||| :  |:||  ||   |  | |||:|       |  |  :|   |   |   ||| |
Db   2947 RAVRASLIAQGGRAAICGRYLFNWAVKTKLKLTPLPEASRLDLSGWFT-----VGAGGGD 3001

Qy   2831 VFITPQRRLQKFJVKYLAVIVPALGLIAVGLAI 2863
           :  |  |||  |  |||| :| || |:|||| |
Db   3002 IYHSVSHARPRLL------LLCLLLSVGVGI 3027
```

RESULT 13
```
Q68801
ID   Q68801;       PRELIMINARY;       PRT;      3019 AA.
AC   Q68801;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96226020; PubMed=8627233;
RA   Tokita H., Okamoto H., Iizuka H., Kishimoto J., Tsuda F.,
RA   Lesmana L.A., Miyakawa Y., Mayumi M.;
RT   "Hepatitis C virus variants from Jakarta, Indonesia classifiable into
RT   novel genotypes in the second (2e and 2f), tenth (10a) and eleventh
RT   (11a) genetic groups.";
RL   J. Gen. Virol. 77:293-301(1996).
DR   EMBL; D63821; BAA09890.1; -.
DR   HSSP; Q8JYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR001917; Aminotrans_II.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
```

```
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3019 AA;  328211 MW;  AF7A6774BC6D95FA CRC64;
```

```
Query Match             21.4%;
Best Local Similarity   30.6%;
Matches  985;     Conservative 411;

Qy     6 TQTSPVP-APRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHDGLQT 64
           :: || ||:  .||:  || ||  | || |   |  |   ::|
Db    60 SRRQPIPRARRTEGRSWAQPGYPWPLYGNEGCGWAGWLLSPRGSRP------------ 105

Qy    65 LAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTHTPLVGAVVRPVCQIV 124
            |||  |    |     :|::|:|||||||||::|||||| |||||| ||||||||
Db   106 ------------SWGPNDPRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPV-GGVARALAHGV 155

Qy   125 RLLEDGVNWATG--WFGVHLFVVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIY 180
            |::||| | ||    |  : |:| |||   |  |  :| | |     |: ||::|::
Db   156 RALEDGINFATGNLPGCSFSIFLLALLSCLLITPAGLEYRNASGLYTV-TNDCSNGSIVY 214

Qy   181 CSPSTCLHEPGCVIC------ADECWVPANP--YISHPSNWTGTDSFLADHIDFVMGALVT 233
            |  | |  | |  |        :: | |        |:  |   :|  |  ||| || |
Db   215 EAGDVILHLPGCIPCVRLNNASKCWTPVSPTVAVSRPGAATAS---LRTHVDMMVGAATL 271

Qy   234 CDALDIGELCGACVLVG---DWLVR-HWLIHIDLNETGTCYLEVPTGIDPGFLG----- 283
            |  |||::| |  |:||    || || : :|: |  |  || | | ::|
Db   272 CSALYVGDLCGALFLVGQGFSWRHRQHWTVQ-DCN--------CSIYPGHLTGHRMAW 320

Qy   284 --FIGWMAAGKVEAVIFLTKLASQVPYAIATMFSSV---HYLAVGALIYYASRGKWYQLLL 338
              ||  |||  | | |    |  |  |  |     |      | | |   ||:|:|| ||
Db   321 DMMMNWSPA-------MTLIVSQVLRLPQTMFDLVIGAHWGVMAGVAYYSMQGNWAKVFL 373

Qy   339 ALMLYI-----------EATSG-----------NPIRV---- 354
            | :  |              |                 :::|
Db   374 VLCLFSGVDASTTITGGVAASGAFTITLSFSTGAKQPLHLVNTNGSWHINRTALNCNDSL 433

Qy   355 --------------PTGC-SIAEFCSPL---------------MIPC 371
                            || | |  | :||                |||
Db   434 NTGFIAGLLYYHKFNSSGCVERMSACSPLDRFAQGWGPLGPANISGPSSEKPYCWHYAPR 493

Qy   372 PCHSYLSENV-SEVICYSPKWTRPITLEYNNSISWYPYT---------IPGAR------- 414
            || :   :::  | | | |    | | | :  |:  | ||        |  :
Db   494 PCDTVPAQSVCGPVYCFTPS---PVVVGATDKRGAPTYTWGENESDVFLLESARPPTEPW 550

Qy   415 -GCMVKFKNNTW----GCCRIRNVPSYCTMG------------------TDAVWNDTRNTYEACG 456
            .||: |||  |     || | || | |:|                    ||  ::|:| || |
Db   551 FGC------TWMNGSGVKTCGAPPCHIYGGREGKSNNSLVCPTDCFRKHPDATYNRCG 603

Qy   457 VTPWLT--------TAWHNGSALKLAILQ---YPGSKEMEKPH-NWMSG-HLYFEGSD 501
            | ||:|         || |   |  | |     | |  |:    |  |  || || |
Db   604 AGPWLTPRCLVDYPYRLWHYPCTVNYTIFKVRMFVGGLEHRFNAACNWTRGERCNLEDRD 663
```

Mon Oct 31. 10:27:45 2005

us-10-009-002-2.rup

```
Qy  1633  LGAVQNEIC---TTHPVTKYIATCMAADLEVATSAWVLLGGVMAALTAY----------  1678
                ||            :  :   |||
Db  1588  RRCWSITSV------PTGATVAPVVDEEI--------------------VEECASFIPL-  1623

Qy  1679  --CLSVGSVVIVGHLVLGGKPALVPDKEVLYQQYDEMEECSRAAPYIEQAGGIAQQFKEK  1736
            :|||        :|  ||  | |  :  ::::       |  |  :  :
Db  1624  MVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATTLAIIEYCCGLVTLPDNPFASCV  1683

Qy  1737  VIGLLQQADQKAADIKPIA-TPYWQKLETFWSKHMWNFVSGIQYLAGLSTLPGNPAIASL  1795
            :  :                 |                    |
Db  1684  FAPIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAGATALGTWTSVGF---  1740

Qy  1796  MAFTASVTSPLITNQTLLFNIMGGWVASNLAPPPASTAFVVSLAGAAVG---SIGLGKV  1852
            :  |                |  |         ::              |||| |
Db  1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVGVGVLSACAMFALTT  1800

Qy  1853  LLDILAGYGAGVAGALVAFKIMGGEMSTEDMVNLLPAILSPGALVVGVI--CAAILRRH  1910
            | |  | | |  | :   |    :     :           |   ||    |  |
Db  1801  AGPD---HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLH  1856

Qy  1911  VGPGEGAVQWMNRLIAFASRGNHVAPTHYVPESDAAAKVTALLSSLTVTQLLRLHQWIN  1970
            | | |         |  |       |     :       |      :::  || | |
Db  1857  ----TPTEDDCGLIAWGLEIWQYVCNFFVICFNVLXAGVQSMVNIPGCPFYSCQKGYKGP  1912

Qy  1971  EDYPTPCDGN------WLYDIWNWVCT--VLADFKLWIGAKIILPKMPGIPFLSCQKGYRGT  2023
            || |||           : | |     :  | | || | | | ||  |  | || | |
Db  1913  WIGGSMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPT-D  1971

Qy  2024  WRGDGVVSTRCPCGALLSGHVKNGTMRLV-GPRWCANTWHGTPINGYTTGPSTPAPSYA  2082
            |  |     || |  ||  |  |   |      |     |    |   | |  | | ||
Db  1972  WTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFTQVP---------------------  2010

Qy  2083  YSRALWRVASDSYVEVRKVGDFHYVTGTTDDGLKCPCQVPLPEFFTELDGVRLHRYAPVC  2142
            :                 |        |   |  |
Db  2011  -PTLRAAVA--------VDGVQVQCYLGEPKTPWTTSACCYGPDGKGKTVKL--PFRVDG  2059

Qy  2143  RPLLRDDVTFTVGLNSYVIGSQLPC---EPE--------------PDVAVVTSMLQDPSHITV  2188
              |   | |    |    | |
Db  2060  HTPGVRMQLNLRDALETND--------------CNSTNNTPSDEAAVSALVFKQELRRTNQL  2107

Qy  2189  ETAKRRLDRGSPPSLASSASQLSAPSRKATCTTHGRHPDAELITANLLWRQEM-------  2242
            |     | || |   |  | |  | |
Db  2108  LEAISAGVDTTKLPAPSIEEVVVRKRQFRARTGSLTLPPPPRSVPGVSCPESLQRSDPLE  2167

Qy  2243  ------GSNITRVES---ESKVVILDSFE-------------------PLRACDDEDELSVAAE  2278
                   |
Db  2168  GPSNLPPSPPVLQLAMPMPLLGAGECN----------PFTAIGCAMTETGGPDDLP  2214

Qy  2279  CFKKPPKYPPAL-----PIWARPDYNPPLVEPWKDPDYVPPTVHGCALP------PQKLP  2327
            |  | |           |
Db  2215  SYPPKK--------------------------EVSEWS---------DESWSTATTASSYVTGPPYPKI  2248

Qy  2328  PVPPPRRKRTIVLSESTVSKALASLAEKSFPQPTCSAEDESTSGVGTQSGSLTGP--VQL  2385
            :: |                            |      |  | | | | |
Db  2249  RGKDS-TQSAPAKRPTKKKLGKSEFS--------------CSMSYTWTD-VISFKTAS  2290

Qy  2386  DDDDSDNESHSSMPPLEGEPGDPDLSSGSWSTVSGEQSVVCCSMSYSWTGALITPCAAE  2445
            |  | | |                          |      |   | |  |
Db  2291  KVLSATRAITSGFLKQRSLVVTEPRDABLRKQKVTINRQPLPPSYHKQVRLAKEKASK  2350

Qy  2446  EEKLPISPLSNSLLRHHNLVYSTSSRSAAQRQKKVTFDRLQVLDDHYNTTLKEIKELASG  2505
            :      |  |     |       |    |
Db  2351  VVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVE--------AGEIP  2404

Qy  2506  VKAELLSVBEEACRLVPSHSARSKF-GYGAKEVRS-LSSKAINHINSVWEDLLEDNTTPIP  2563
            :               |   |
Db  2405  SHYRQTVIVPKEEVFVKTPQKPFKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAVMGDA  2464

Qy  2564  -----TTIMAKNEVFAVAPHKCGGRKPARLIVYPDLGVRICEKRALYDVIQKLPSAIMGSA  2618
Db  2465  YGF-VDPRTRVKRLLSMWS---PDAVGATCDTVCFPDSTITPEDIMVETDIYSAAKLSDQHR  2521
```

```
Qy   502  TPIVY--FYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQGFYSDVKDLIATGLITKDKA   559
            : ||                              ||||
Db   664  RSEMYPLLHSTEQAILP-----------CSFVPIPA-------LSTGLIHLHQN   700

Qy   560  WKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAYLCYLSLCFGRASGYPLRPV   619
            :      :     |          |  ||     |      |         |  :
Db   701  IVDVQVIYGISSGLVGWAIKWEFVILIFLLLADARVCCVLWMMML--------------   745

Qy   620  LPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAALLGFVMAAGLPLTFVAAA   679
            |: :  :|  ||| |||  |   |    |    |   | :
Db   746  ------------------ISQAEAALENLI   763

Qy   680  AAQPDYDWWVRLLVAGLVLWAGRNRGHRIAL--LVGPWPLVALLLTLLHLVTPASAF----   733
            ||            |||        | |     |  | | | | | |   | |
Db   764  AAGTHGIWW--SLVAFCVAWHVRGRIFPIAVYSIVGLWPLL----LLVLMLPYRAYAWTG   817

Qy   734  -DTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQRWEN----WFWNVTLRPE   788
             |                    |  |              | |            ||
Db   818  TDTSTLGA-GVLSLFALFTLS--PWYKHWIARLIWWNQYTIARCEAALQIWVPPLLARGA   874

Qy   789  RFFLVLVCFPGATYDALVTFCVCHVALLCLTS---SAASFFGTDSRVRAHRMLVRL----   841
            | |  |    |  :    |    | |||             || |     : |
Db   875  RDGIILLA--GLFYPALV-FDITKLLLLAILGPLYIIQASLVRVPYFVRAH-AVVRLCILV   930

Qy   842  -----GKCHAWYSHYVLKFPLLVFGENGVFFYKHLHGDVLPNDFASK-----LPLQEP--   889
                    | | |    | :||           |  |        :    |  |
Db   931  -----YVQMVLALARGFNTYLYDHLSP---MTDWAAEGLKDLAVAVEPVI   980

Qy   890  FFPFEGKARVYRNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPD------GWAITAPFT   944
            | | |||                |   |     || :                 |
Db   981  FSPMEVKVITWGAD--TTACGDILCGLPVSARLGKEVLLG---PADDYRSMGWRLLAPIT  1035

Qy   945  LQCLSRGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRL  1004
            |  |       |         |      |             |
Db  1036  AHAQQTRGLFGTIVTSLTGRDKNIVTGEIQVLSTSTQTFLGTSVGGVMWTVVHGAGSRTL  1095

Qy  1005  AHPTGSIHPITVDAAN-DQDI--YQPPCGAGSLTRCSCGETKGYLVTRLGGSLVEVNKSDD  1061
            |  | :   | :             ||           |          | | |
Db  1096  A----GNKRPALQMYTNVDQDLVGWPSPPGAKSLVPCTCGSADLYLITRDADVLPARRRGD  1152

Qy  1062  PYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQY  1121
              |  || |  |     | ||    | |        |     |
Db  1153  STASLLSPRPLACLKGSSGGPIMCPSGHVAGIFRAAVCTRGVAKALQFIPVESLSAQTRS  1212

Qy  1122  TAHATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKY  1181
            ||                |  ||||| ||             | | ||| || |
Db  1213  PSFSDNSTPPAVPQTFQVGYLHAPTGSGESTKVPASYVAQGYTVLVLNPSVAATLGFGRF  1272

Qy  1182  MHATYGVNPNCYFNGKCTNTGASLTYSTYGMYLT-GACSRN-YDVIICDECHATDATTVL  1239
                |||     | | |    | |||| |      |  |     |  | || |
Db  1273  MSHAYGIDPNVRTGTRTITTGAKLTYSTYGKFLADGGCSGGAYDVIICDECHAQDATSIL  1332

Qy  1240  GIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKK  1299
            ||  |   |   |  | | ||||||||  |       | |  | | | |     |  :|
Db  1333  GIGTVLDQAETAGARGLVVLATATPPGSITVPHSNIEEVALITGEGEIPFYGRAIPLGVIKG  1392

Qy  1300  GRHLIFEATKGHCDELANELARKGITAVSYYRGCDISKIP-EGDCVVVATDALCTGYTGD  1358
            |||||       | | | | ||    | |||||    | |    ||  |||  | |||
Db  1393  GRHLIFCHSKKKCDELAAKQLTSLGVNAVAFYRGLDVSVIPTQGDVVVCATDALITGYTGD  1452

Qy  1359  FDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCT  1418
            ||||||   ||  ||   |    :|  | |   | |  | | | | | |  ||  ||
Db  1453  FDSVIDCNVAVEQYVDFSLDPFFSIETHTVPQDAVSRVSQRRGRTGRGKSGTYRYVSPGER  1512

Qy  1419  PSGMVPECNIVEAPDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFS-MVN  1477
            ||| || :   | |      | |  |        |  |  |
Db  1513  PSGMFDSVVLCECYDAGCAWYELTPSETTVRLRAYLSTPGLPVCQDHLEFWEGVFTGLTH  1572

Qy  1478  PEPSFVNTAKRTADNVVLLTRAQLQLCHQYGYAAPNDAPRWQG-ARLGKK---PCGVLWR  1533
            |  :   ||                  |   |
Db  1573  IDAHFLSQTKQQGLNFPYLTAYQATVCARAAALPPSWDETWKCLIRLKPTLHGPTPLLYR  1632

Qy  1534  LDGA--DACPGPEPSEVTRY-QMCFT--EVNTSGTAALAVGVGVAMAYLAIDTFGATCV  1587
```

```
Db  2619 YGFQYSPKQRVEYLLKKMWNSKKTPLGFSYDTRCFDSTVTEQDIRVEESIYQACDLKDEAR 2678
              : .: .||    |||  |::|| |.: ||:|
Qy  2522 AGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMKN  2581
Db  2679 RVITSLTERLYCGGPMFNSKGQHCGYRRCRASGVLPTSFGNTVTCYLKAKAATKAAGIKD  2738
Qy  2582 PRFLICGDDCTVTWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSNV  2641
              |.||.|||| :| : |:|.| :: |||: : || ::    :  |:  :||:||.||||
Db  2739 PSFLVCGDDLVVIAESAGIDEDKSALRAFTEAMTRYSAPPGDPPQPTYDLELITSCSSNV  2798
Qy  2642 TSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHFM  2701
             : .|   |:||:|||| :.||.: |.| |  | :||. |:: :||:|| :| :::.||:
Db  2799 SVAHDGAGKRYYLTRDPETPLARAAWETARHTPVNSWLGNIIMYAPTIWVRMVIMTHFF   2858
Qy  2702 EQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEARSVVRXTNAEILRVSQSLT  2761
              | : |:| || .:|:  |.:   |::|||:||| |||:||  |::::   ||||:|:||
Db  2859 SILQAEQELEKALDFEMYGAVVSVTPLDLPAIIERLHGLSAFSLHSYSPVELNRVAGALR  2918
Qy  2762 DMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWH------ATSRPLPDLDKTSVAR  2815
              .:||||||||::|| ||:::|:|||.|||:  |:|:        |||||| :|:|||
Db  2919 KLGIPPLRAWRHRARAVRAKLISQGG-KAKICGLYLFNWAVRTKAKLTPLPQAGLLDLSR  2977
Qy  2816 YTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGL                    2856
              . | .. |.:|. |::|:: : :| |:: ::.|:: . |
Db  2978 WFT-----VGAGNDIYHSVSRARSRHLLGLLLLTVGVGI                      3013
```

```
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; Viral_RdRP; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3033 AA;  330477 MW;  3616B01905C3143E CRC64;

Query Match              21.4%;  Score 3289.5;  DB 2;  Length 3033;
Best Local Similarity   29.9%;  Pred. No. 1.4e-204;
Matches  985;  Conservative 444;  Mismatches 1155;  Indels 713;  Gaps 110;

Qy    8 TSPVPAPRTRKNKQTQASYPVSIK---------------TSVERGQRAKRKVQRDARPR  51
          |.: :.| |.|::.     . |:||               |.:   |.:|:| |||:||
Db    3 TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR  59
Qy   52 NYKIAGIHDGLQTLAQA------------ALPAHG---------------WGRQDPRHKSRN   86
          . :. | |.:|..|:              ||. |                |||.||||.||
Db   60 GRR-QPIPKDRRSVGKSWGKPGYPWPLYGNEGCGWTGWLLSPRGSRPTWGPTDPRHRSRN  118
Qy   87 LGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG--WFGVHLF  143
          ||.: | . :| || :|:: ||:| ||. . | : .:::::||||.||||   :| :.:|
Db  119 LGKVIDITTCGFADLMGYIPVVGAPV-GGVARALAHGVRVLEDGINYATGNLPGCSFSIF  177
Qy  144 VVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVICADE---  198
          :. .:. ||| : |.:   |:|:| : .:|.:.. .:  |  :.| || |: |.::
Db  178 LLALLSCVTVPVSAVEVRN-ISSSYYATNDCSNNSITWQLSNAVLHLPGCVPCENDNGTL  236
Qy  199 -CWVPANPYISHPSNWTGTDSFPLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV---  254
           ||:   . :.||.||||.||: .| .:| :||| || ||  :|.: |.| :::.:|:
Db  237 RCWIQVTPNVAVKHRGALTHN-LRTHVDMIVMAATVCSALYVGDVCGAVMIVSQALIVSP  295
Qy  255 -RHWLIHIDLNETGTCYLEVPTGIDPGPLGFIGWMAGKVEAVIFLTKLASQVPYAIATMF  313
           | .:|  .:.|. .| ||   |::.| :.|::    .| |:: |.: |  |||:: .:
Db  296 ERH-----NFTQECNCSIYQGHVTGQRMAWDMMLNWSPTLTMILAYAARVPEMVLEII    348
Qy  314 SSVHYLAVGALIYYASRGKWYQLLLAMLY-IEAT------------                347
           .. : .|||: : :: |. :.:: |:| ||.|
Db  349 FGGHWGVVFGLAYFSMQGAWAKVIAILLLVAGVDATTYSSGAQVGRALGGLTGIFSPGAK  408
Qy  348 --------------------------------SGNPIRVPTGCSIAEF---            363
                                           :|:.:..|.|| .  :
Db  409 QNIQLINTGSWHINRTALNCNDSLQTGFIASLFYTHRFNSSGCPERLSSCRGLDDFRIG   468
Qy  364 --------------C--SPLMIPCPCHSYLSENV-SEVICYSPKWTRPITLEYNN       401
                        |   |.:: |  |.   ..|   :.||.|.| .|  :.|.  |
Db  469 WGTLEYETNVTNDEDMRPYCWHYP--PKPCGIVSAKTVCGPVYCFTPS--PVVGTTD    522
Qy  402 ------SISWYPYTIPGARGCMVKFKNNT--------WGC--            CRIR--   431
                . ::.| .   . || :.|.|::.        |||             ||:|
Db  523 RQGVPTYSW-----GENETDVFLLNSTRPPQGAWFGCTWMNGTGFTKTCGAPPCRIRRD  576
Qy  432 -NVPSYCTMGTDAVWNDTRNTYEACGVTPWLT------TAWHNGSALKLAILQ---Y    478
           |.:| |.:| || | |.| .: | | . ||.        |:|.|: ::|.:|   :
Db  577 YNSTLDLLCPTDCFRKHPDATYLKCGAGPWLTPRCLVEPYRLMWHYPCTVNFTIFKVRMY  636
Qy  479 PGSKEMFKPHNWMSGHLYF---------------EGSDTPIVFYFYDPVNSTLLPPERWARLPG  526
          .|.  :|  .|.::|.|::                .|.  : ::|.|:  .||||
Db  637 VGGVE-----HRLSAACNFTRGDRCRLEDRDRGQQSPLLH-------ST---TEWAVLPC   681
Qy  527 TPFVVRGSWLQVPGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLi   586
          | :.:||| | | ||:||||||.|||:|||||||||||| |:||||| ||||||||:|
Db  682 S-------FSDLPALSTGLLHLHQNIVDVQYLYGLSPALTRYIVKWEWVVLL          726
```

```
RESULT 14
Q7T7H7

ID   Q7T7H7      PRELIMINARY;      PRT;   3033 AA.
AC   Q7T7H7;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tanabe Y., Nagayama K., Enomoto N., Izumi N., Tazawa J., Sato C.,
RA   Watanabe M.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY232745; AAP55700.1; -.
DR   HSSP; Q8JYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
```

```
Qy   1582 FGATCVRRCWSITSVPTGATVAPVVDEE------EIVEECASPIPL-EAMVAAIDKLKST 1634
               ||:  | | |   |                  || : :   ||
Db   1682 GCVSIIGRLHLNDQV-----VVAP--DKEILYEAFDEMEECASKAALIEEGQRMAEMLKSK 1735

Qy   1635 I------TTTSPFTLETALEKLNTFLGPHAATTLAIIEYCCGLVTLPDNPFASCV 1683
          |       :  ||  : | :: |   :: |   |  | |  :: ||   |: :
Db   1736 IQGLLQQATRQAQDVQP-VIQSSWPKLEQFWAKHMWNFISGIQYLAGLSTLPGNPAVASM 1794

Qy   1684 FAPIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVG---F 1740
           | : | |: |    : |||| :|| ||:: :|   ||| |||||||| |        |
Db   1795 MAFSAALTSPLPTSTTILLNIMGGWLASQIAPPAGATGFVVSGLVGAAVG---SIGLGKI 1851

Qy   1741 VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVVSAFNPAAGVVGVLSACAMFALTT 1800
          : | :|||| | |: | :||   ||::| :: |||| |:| || ||||:||| | :
Db   1852 LVDVLAGYGAGISGALVAFKIMSGEKPSVEDVVNLLPAILSPGALVVGVICAAILRRHIG 1911

Qy   1801 AGPD--HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTP 1858
           | |   |  |: :: |  ::  :  :| : | ||    :  ::|:     |: |  |:
Db   1912 QGEGAVQWMNRLIAFASRGNHVAPTHYVAESDASQRVTQVLSSLT----ITSLLRRLHAW 1967

Qy   1859 TEDDCGLI---AWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFVSCQKGYKGPWIG 1915
          :| |  |     |||: | :| : |  :|| ::|  |::    |:| ||||||| |:| |
Db   1968 ITEDCPVPCSGSWLQDIWDWVCSILTDFKNWLSS--KLLPKLPGLPFISCQKGYRGVWAG 2025

Qy   1916 SGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARP-DPTDWTS 1974
           | :  ||||| | | ::   || |  |  ||| :: : | ::| ::|   | ||  ::
Db   2026 TGVMTTRCPCGANISGHVRMGTMKI-TGPKTCLNLWQGTFPINCYTEGPCVPKPPPNYKT 2084

Qy   1975 LVNYGVRDYCKYEKMGDHIFVTAVSSPNVCFTQVP---PTLRAAVAVDGVQVQCYLGEP 2031
          ::  | |  |    : | ||  | :||||    |     | |  || |||:::  : | |
Db   2085 AIWRVAASEYVEVTQHGSFSVYTGLTSDNL--KVPCQVPAPEFFSWVDGVQIHRFAPTP 2141

Qy   2032 KTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSD--- 2088
               |   |  :    |  |  |: : |   |  :    |  | | : :
Db   2142 -------GPFFRD---EVTFTVGLNSFVVGSQLPCDPEPDTEVLASMLTDPSHITA 2187

Qy   2089 EAAVSALVFKQELRRTNQLLEAISA----------GVDTTKLPA- 2122
           ||  |:    |  :: :   : |            |||: | |
Db   2188 EAAARRLARGSPPSQASSSASQLSAPSLKATCTTHKTAYDCDMVDANLFMGGDVTRIESD 2247

Qy   2123 ------PSI-EBVVVRKRQFRARTGSLTLPPPPRSVPGVSCPESLQ 2161
                 || : |:|  |:|  | |
Db   2248 SKVIVLDSLDSMTEVEDDREPSVPSEYLIRRKF-------------- 2281

Qy   2162 RSDPLEGPSNLPP----------SPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGPDDLPS 2215
             |||  ||  ||           || :|  :   :::        |  | |    |     | |
Db   2282 ------PPALPPWARPDYNPPVIE-TWKRP----DYEPPTVLGCALPST----PQVPV 2324

Qy   2216 YPPKKE--------------------VSEWSDESWSTA----- 2237
          ||| :|                    |:| | : |:
Db   2325 PPPRRRARVLTQDNVEGVLREMADKVLSPLQDHNDSGHSTGADTGGSVQQPSDETAAS 2384

Qy   2238 SYVTGPPYPKIRGK--------DSTQSAPAKR----------PTKKKLGKSEFSCGSM 2276
          |:  || |||  |          ||:| || |          || |::| | : |||
Db   2385 ETGSLSSMPPLEGEPGDPDLEFEPARSAPPSEGECEVIDSDSKSWSTVSDQEDSVICCSM 2444

Qy   2277 SYTWTD-VISFKTASKVLSATRAITSGFLKQRSLVVVTEPRDAELRKQKVTINRQPLFPP 2335
          || ||  |||| |||||:| | |:|| | |:| | |:: |::| |: |:|| :|:|| |
Db   2445 SYSWTGALITPCCGPEEEKLPINPLSNSLMRFHNKVYSTTSRSASLRAKKVTFDRVQVLDT 2504

Qy   2336 SYHKQVRLAKEKASKVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQ 2395
          || :::::|:: |||| | | :  | || |||:|||  |  : ::  :|:||: :  :
Db   2505 HYDSVLQDVKRAASKVSARLLSIEEACALTPHSAKSRY-GFGAKEVRS-LSRRAVNHIR 2562

Qy   2396 KCVEAGEIPSHYR-QTVIVPKEEVFVKTPQKPKKKPPRLISYPHLEMRCVEKMYYGQVAP 2454
           | |   |  ::| : :  :|:|||    |  | ||| ::  |||  |||||:  : ||||
Db   2563 SVWEDLLEDQHTPIDTTIMAKNEVFCVDPTKGGKKPARLIVVPDLGVRVCEKMALYDIAQ 2622

Qy   2455 DVVKAVMGDAYGF-VDPRTRVKRLLSMW--SPDAVGATCDTVCFDSTITPEDIMVETDIY 2511
           :|||:||| |||  ||  ||  ||:  |   ||||  || ||||||:||  || : | ||
Db   2623 KLPKAIMGPSYGFQYSPAERVDFLLKAWRSKKDPMGFSYDTRCFDSTVTERDIRTEESIY 2682

Qy   2512 SAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSNSLTCWLKVN 2571
```

```
Qy    587 LLGLCGSKYLILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFF 646
           |: |   ||   |  ||:| ||:  ||
Db    727 FLLLADAR--VCACLWMLI----------------------ILGQAEAALEKLII- 757

Qy    647 ICCYLRCRLRYAALLGFVPMAAGLPLTFFVAAAAQPDYDWVRLLVAGLVLWAGNRGH 706
                         |  :|     || :::
Db    758 ----------LHSASASANGPLWFFIFFTAA--------WYLKGRVVPVATYS-- 793

Qy    707 RIALLVGPWPLVALLTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGF---FAHLLP 763
           :: || :   :  ||  | | |::| :  | :: ||  : ::::  |         |
Db    794 ----VLGLWSF--LLIVLALPQQAYAMDAABQGELGL---VILVIISIFTLTPAYKILLS 844

Qy    764 RCALVNSYLW--------QRW-------ENWFW-NVTLRPERFFLV---LVCFPGATY 802
          | |:|||:|         ||          :   |:  :|  |::       |  | |
Db    845 RSVWWLSYMLVLABAQIQQWVPPLEARGRGDGIIWVAVILHPRLVFEVTKWLLAILGPAY 904

Qy    803 DALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRM----LVRLGKCHAWYSHYVLKFFL 857
           ||:|  | |||||  |         |:|||||   :   ||||| | || ||||:|||
Db    905 ----------LLKASLLRVPYF-----VRAHALLRVCTLVR----HLAGARYV-QMLL 942

Qy    858 LVFGE-NGVFFYKHLHGDVLPNDFASKLPLQ----------EP-FFPFEGKARVY 900
          |  |   | ::    || :  ::||::|  |             | : :| | : :
Db    943 LTMGRWTGTYIYDHLS----------PLSTWAAQGLRDLAVAVEPVVFSPMEKKVIVW 990

Qy    901 RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPD-----GWAITAPFTLQCLSERGTLS 955
           : | :||||||::| |||:|||| :|  || :        ||   || :: | | |||
Db    991 GAE--TVACGDILHGLPVSARLGREVLLG----PADSYTSKGWKLLAPITAYTQQTRGLLG 1045

Qy    956 AMAVMTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHGSKGRRLAHPTGSIHPIT 1015
          |:|::|||| :     ::|  |:  ||:|  |::||:| ||: |: |  |:|| :|  |
Db   1046 AIVVSLTGRDKNEQAGQVQVLSSVTQSFLGTCISGVLWTVVTHGAGNKTLAGPKGPVTQMY 1105

Qy   1016 VDAANDQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVA 1075
            :|  | :: || |:|| || ||| | ||||:| | : | :: :: |:::|   | ||
Db   1106 TSAEGDLVGWSPPGTKSLDPCTCGAVDLYLVSRNADVIPVRRKDDRRGALLSPRPLSTL 1165

Qy   1076 KGSSGAPILCSSGHVIGMFTA---ARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPT 1132
          ||||| || ||:||| | |         :|  | || | |||: |: | || |:| ::| |
Db   1166 KGSSGGPVLCSRGHVVGLFRAAVCARGVAKSIDFIPVESLNIATRTPSFSDNS---TPPA 1222

Qy   1133 VPNEYSVQILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNC 1192
          || :| | :| ||||||||| | : :: : ||||||||||| | : :: |::| |||||
Db   1223 VPQTYQVGYLHAPTGSGKSTKVPAAYASQGYKVLGLNPSVAATLGFGAYMSKAHGINPNI 1282

Qy   1193 YFNGKCTNTGASLTYSTYGMYLT-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPS 1250
          : |   :||  | |||||| ::  | |   ||:|||||||| ||||:|| ||:|: :| :
Db   1283 RTGVRTVTTGDPITTYSTYGKFVADGGCQAGAYDIIICDECHSVDATTILGIGTVLDQAET 1342

Qy   1251 KNVRLVVLATATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKK 1310
           ||||||||||||||| | | |:|::  | |:| ||: | || | || |||||| :|||
Db   1343 AGVRLVVLATATPPGVTTPHSNIEEVALGHEGEIPFYGKAIPLAFIKGGRHLIFCHSKK 1402

Qy   1311 HCDELANELARKGITAVSYRGCDISKIP-EGDCVVVATDALCTGYTGDFDSVYDCSLMV 1369
           |||| ::|||:|:  | |||  :| || |||||||||||| |||||||||||:||::|
Db   1403 KCDELAAALRGMGVNAVAYYRGLDVSVIPTQGDVVVVATDALMTGYTGDFDSVIDCNVAV 1462

Qy   1370 EGTCHVDLDPTFMGVRVCGVSAIVKGQRRGRTGRGRAGIYYVYDGSCTPSGMVPECNIV 1429
           : :: ||||||:     : :| :: ||||||||||| |:| |  |  ||||: ::|::
Db   1463 TQIVDFSLDPTFTITTQTVPDAVSRSQRRGRTGRGRLGTYRYVSSGERPSGMFDSVVLC 1522

Qy   1430 EAPDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKR 1488
          | :|| :|||| : :   |||:|: | ||||  | :|:||: :|   ||| |:| ::|:
Db   1523 ECYDAGAAWYELTPAETTVRLRAYFNTPGLPVCQDHLEFWEAVFTGLTHIDAHFLSQTKQ 1582

Qy   1489 TADNVLLTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADAC-------- 1540
           | : : ||| : :| |::|: |          :| || |:  :  :  :
Db   1583 GGENFAYLTAYQATVCAR---AKAPPPSWD----VMWK------CLTRLKPTL 1622

Qy   1541 PGPEP--------------SEVTRY----QMCFTEVNTSGTAALAVGVGVAMAYLAIDT 1581
           ||                :|||   :|   ::| :| :|  || |  :|:: || |
Db   1623 TGPTPLLYRLGAVTNEVTLTHPVTKYIATCMQADLEIMTS-TWVLAGGVLAAVAAYCLAT 1681
```

```
Db        2683 QACSLPQEARTVIHSLTERLYVGGPMTNSKGQSCGYRRCRASGVFTTSMGNTMTCYIKAL 2742
               |||    |:  | ::   |: ||||       || :||||:||     |||||||::|:|
Qy        2572 AAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSL 2631

Db        2743 AACKAAGIVDPIMLVCGDDLVVISESQGNEEDERNLRAFTEAMTRYSAPPGDLPRPEYDL 2802
               |||:|| |:| |:|:||||| |||:| :: ::: ||  ||  ||||:|  ||:|||||||
Qy        2632 EELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLW 2691

Db        2803 ELITSCSSNVSVALDPRGRRRRYFLTRDPTTPITRAAWETVRHSPVNSWLGNIIQYAPTIW 2862
               |:|||||||| |||| |:: ||||||:| |:  : || :|||||| :: || ||||::
Qy        2692 VSRVLAVHFMEQMLPEDKLPETVTFDWYGKNTVPVEDLPSIIAGVHGIEARSVVRVRTNA 2751

Db        2863 VRMVIMTHFFSILLAQDTLNQNLNFEMYGAVYSVNPLDLPAIIERLHGLDAFSLHTYSPH 2922
               ||   :|||| :|  : | :  |:|:|  || | : |:|  |:||::||| ||| |: |
Qy        2752 EILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WHATSR----PLP 2806

Db        2923 ELSRVAATLRKLGAPPLRAWKSRARAVRASLIAQGGRAAICGRYLFNWAVKTKLKLTPLP 2982
               |||::   ||||||||||||:|||||||||| |:||| |:|||:|   |  ||||| ||
Qy        2807 DLDKTSVARYTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAI 2863

Db        2983 EASRLDLSGWFT-------VGAGGDIFHSVSRARPRLL-------LLCLLLLSVGVGI 3027
               |:| |::|||     ::::||||         |||           :||  :|:|||
Qy        2983 SQLCSSNVSVALDPRGRRRRYFLTRDPTTPITRAAWETVRHSPVNSWLGNIIQYAPTIW 2983
```

```
RESULT 15
Q7T7I7;

ID   Q7T7I7       PRELIMINARY;   PRT;   3033 AA.
AC   Q7T7I7;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Polyprotein.
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tanabe Y., Nagayama K., Enomoto N., Izumi N., Tazawa J., Sato C.,
RA   Watanabe M.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY233735; AAP55690.1; -.
DR   HSSP; Q8JYS1; 1CWX.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   GO; GO:0019087; P:viral transformation; IEA.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001410; DEAD/DEAH_N.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR002522; HCV_capsid.
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   InterPro; IPR000745; HCV_NS4a.
DR   InterPro; IPR001490; HCV_NS4b.
DR   InterPro; IPR002868; HCV_NS5a.
DR   InterPro; IPR002166; HCV_RdRP.
DR   InterPro; IPR004109; Peptidase_S29.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR002518; Pept_U39_HCV_NS2.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF01543; HCV_capsid; 1.
```

```
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
DR   Pfam; PF01538; HCV_NS2; 1.
DR   Pfam; PF02907; HCV_NS3; 1.
DR   Pfam; PF01006; HCV_NS4a; 1.
DR   Pfam; PF01001; HCV_NS4b; 1.
DR   Pfam; PF01506; HCV_NS5a; 1.
DR   Pfam; PF00998; Viral_RdRp; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   PROSITE; PS00190; CYTOCHROME C; UNKNOWN_1.
KW   Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW   Polyprotein; Transmembrane.
SQ   SEQUENCE   3033 AA;  330352 MW;  A83492790E9172C0 CRC64;

Query Match             21.4%;  Score 3289.5;  DB 2;  Length 3033;
Best Local Similarity   30.2%;  Pred. No. 1.4e-204;
Matches  985;  Conservative 446;  Mismatches 1189;  Indels 641;  Gaps 109;
```

```
Qy           8 TSPVPAPRTRKNKQTQASYPVSIK-----------------------TSVERGQRAKRKVQRDARPR 51
               |  |||  ||||  |:           ::                      |:  |  ||: |  | :
Db           3 TNPKPQRKTKRNTNRR--PQDVKFPGGGQIVGGVYLLPRRGRPLGVRATRKTSERSQPR 59

Qy          52 NYKIAGIHDGLQTLAQAALPAH--------------------------------------GWGRQDPRHKSRNL 87
               |  ||  || |||   | | |                                        || |  |||||||
Db          60 GRQPIPKDRRSTGKSWGKGPYMPLYGNEGCWAGWLLSPRGSRPTWGPTDPRHRSRNL 119

Qy          88 GILLDYPLGWIGDVTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLFV 144
               ||:| :|  ||| || |  |  |:|| :: |||   |:|| |  |: |   :  ||:||
Db         120 GKVIDTITCGFADLMGYIPVIGAPV-GGVARALAHGVRVLEDGVNYATGNLPGCSFSIFL 178

Qy         145 VCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVICADE----- 198
               ||||   ||||| |   |   |     :|||:|||                          ::
Db         179 LAILLSCVTVPVSAVEIRN-ISSSYATNDCSNNSITWQLTNAVLHLPGGCVPCENDNGTLR 237

Qy         199 CWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV---- 254
               ||  |||  ||  |||       :  |  | : |||   :      |       ||
Db         238 CWIQVTPNVAVKHRGALTHN-LRTHVDMIVMAATVCSALYVGDVCGAVMIVSQALIVSPQ 296

Qy         255 RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTKLASQVPYAIATMFS 314
               |     :      |:            ::  |  |  : || :   ::  :::   :
Db         297 RH------NFTQECNCSIYQGHITGHRMAWDMLNWSPTLTMILAYARVPELALELVF 349

Qy         315 SVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTCGSIAEF---CSPLMPC 371
               |  | || ||  |||:||||  ||:||  |:|:  |||| |: ||          ::|||
Db         350 GGHWGVVFGLAYFSMQGAWAK-VIAILLVAGVDAN--TYSTGAQAGHTISGLSGLFAPG 406

Qy         372 P-----------------CHSYLSENVSEVICYSPKWT-----------RPI----- 395
               |                 |                 |            |||
Db         407 PRQRIHLINTNGSWHINRTALNCNDSLKTGFIASLFYTNNFNSSGCPERLSSCRALDDFR 466

Qy         396 -----TLEYNNSIS--------W-YP-----------YTIPGARGCM------------ 417
                    |   ||:                   :                  :
Db         467 IGWGTLEYETNVINDEDMRPYCWHYPPKPCGIVSARTVCGPVYCFTPSPVVVGTTDRQGV 526

Qy         418 -----------------------VKFKNNT-----WGC---------CRIR---NVPSYCTMG 440
                                      :|||  |     | |         |||:   |  |  |||
Db         527 PTYSWGENETDVFLLNSTRPPQGAWFGCTWMNGTGFTKTCGAPPCRIRRDYNSTLDLLCP 586

Qy         441 TDAVWNDTRNTYEACGVTPWLT------------TAWHNGSALKLAILQ--YPGSKEMFKPH 488
               | || |:| ||:|  |  |||             |:|| | |:| ||     ||  :|| |
Db         587 TDCFRKHPGTTYIKCGAGPWLTPKCLVEYPYRLWHYPCTVNFTIFKVRMYVGGVE----- 641

Qy         489 NWMSGHLYF----------EGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWL 536
               |  |  |:            :::|| :||:||            |      |:|||
Db         642 HRLSAACNFTRGDRCRLEDRDRGQQSPLLH--------ST----TEWAVLPCS----- 682

Qy         537 QVPQGFYSDVKDLATGLITKDKAWKNYQVLYATGALSLTGVTTKAVVLLLGLCGSKYL 596
               |  | ||  | ||  |:   : ::|  | ||  |   |||::| || :  | |||:|
Db         683 ------FSDLPALSTGLLHLHQNIVDVQVLYGLSPAVTRYIVKWEWVVLLFLLLADAR-- 734

Qy         597 ILAYLCYLSLCFGRASGYLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLR 656
               : :| ||          ||||:||||| |  | |:|    |||| |   |    | | |
```

```
Qy  1658 AATILAIIEYCCGLVTLPDNPPASCVFAFIAGITTTPLPHKIMPLSLFGGAIASKLTDAR 1717
Db  1769 MWNFISGIQYLAGLSTLPGNPAVASNMAFSAALTSPLPTSTTILLNIMGGWLASQIAPPA 1828

Qy  1718 GALAFMAGAGTALGTWTSVG---FVFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAG  1774
Db  1829 GATGFVVSGLVGAAVG---SIGLGKIIVDVLAGYGAGISGALVAFKIMSGERPSVEDVVN 1885

Qy  1775 LVYSAFNPAAGVVGVLSACAMFALTTAGPD-HWPNRLLTMLARSNTVCNEYFIATRDIR  1832
Db  1886 LLPAILSPGALVVGVICAAILRRHVGGGEGAVQWMNRLIIAFASRGNHVAPTHYVAESDAS 1945

Qy  1833 RKILGILEASTPWSVISACIRWLHTPTEDDCGLI---AWGLEIWQYYCNFPVICFNVLKA 1889
Db  1946 QRVMQVLSSLT---ITSLLRRLHTWITEDCPVPCSGSWLRDIWDWVCSILTDFKNWLSS  2001

Qy  1890 GVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSN 1949
Db  2002 --KLLPKMPGLPFLSCQKGYRGVWAGTGVMTTRCPCGANISGHVRMGTMKI-TGPKTCLN 2058

Qy  1950 YWRGAVPVNARLCGSARPDPT-DWTSLVVNYGVRDICYKBKMGDHIPVTAVSSPNVCFTQ 2008
Db  2059 LWQGTFPINCYTEGPCVPKPSPNYKTAIWRVAASEVEVTQHGSFSVVTGLTSDNL---K  2115

Qy  2009 VP---PTLRAAVAVDGVQVQCVLGEPKTPWTTSACCYGP--DGKGKTVKL-PFRVDGHT  2061
Db  2116 VPCQVPAPEFFSWVDGVQIHRFAPIP---------GPFFRDEVTFTVGLNSFVVGSHL   2164

Qy  2062 P-GVRMQLNLRDALETNDCNST---------NNTPSDEAAVSALVFKQELRRT------  2104
Db  2165 PWDPEPDTEVLASMLTDPSHITAEAAARRLARGSPPSQASSSASQLSKATCTTHKM     2224

Qy  2105 ---NQLLEA-ISAGVDTTKLPA------------------PSI-EEVVRKRQFRAR     2138
Db  2225 AYDCDMVDANLFMGGDVTRIESDSKVVILDSLDSMTEVEDDREPSVPSEYLIRRRKF--  2281

Qy  2139 TGSLTLPP--PPRSVPGVSCP--ESLQRSDPLEGPS----NLPPSPPVLQLAMP----  2184
Db  2282 ------PPALPPWARPDYNPPVLETWKRPD-YEPPTVLGCALPPTP---QTPVPPRRRR 2331

Qy  2185 ------------------------MPLLGAGECNPFTAIGCA----------        2207
Db  2332 AKVLTQDNVEGVLREMADKVLSPLQDHNDSGHSTGVDTGGDSVQQPSDETPASETGSLSS 2391

Qy  2208 ---------GGPD----DLPSVPPKK---EVSEWSDESWSTATTASSVVTGPPYPKIRGKD 2252
Db  2392 MPPLEGEPGDPDLEFEPAGSAPPSEGECEIVDSDSKSWSTVSDQED---------      2437

Qy  2253 STQSAPAKRPTKKKLGKSEFSCSMSYTWTD-VISFKTASKVLSATRAITSGFLKQRSLVY 2311
Db  2438 ---------SVICCSMSYSWTGALITPCGPEEEKLPINPLSNSLMRFHNKVY         2480

Qy  2312 VTEPRDAELRKQKVTINRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAK 2371
Db  2481 STTSRSATLRAKKVTFDRVQVLDAHYDSVLQDVKRAASKVSARLLSIEEACALTPPHSAK 2540

Qy  2372 SHITGLRGTDVRSGAARKAVLDLQKCVEAGEIPSHYR-QTVIVPKEEVFVKTPQKPTKKP 2430
Db  2541 SRY-GFGAKEVRS-LSRRAVNHIRSVWEDLLEDQHTPIDTTIMAKNEVFCLDPTKGGGKP 2598

Qy  2431 PRLISYPHLEMRCVEKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMW--SPDAVG 2487
Db  2599 ARLIVYPDLGVRVCEKMALYDIAQKLPKAVMGPSQGFQYSPAERVDFLLKAWGSKKDPMG 2658

Qy  2488 ATCDTVCFDSTITPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGY 2547
Db  2659 PSYDTRCFDSTVTERDIRTEESIYLACSLPQEAKTAIHSLTERLYVGGPMTNSKGQSCGY 2718

Qy  2548 RRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAM 2607
Db  2719 RRCRASGVFTTTSMGNTMTCYIKALAACKAAGIVDPIMLVCGDDLVVISESQGNEEDERNL 2778
```

```
Db   735 VCACLWMLI------------------ILGQAEAAALEKLII-----------        757
Qy   657 YAALLGFVPMAAGLPLTFFVAAAAQPDYDWVRLLVAGLVLWAGRNRGHRIALLVGPWP   716
Db   758 ---LHSASASANGPLWFFIFFTAA------WYLKGRVVPAATYS--------VLGLWS   799

Qy   717 LVALLTLLHLVTPASAFDTEIIGGLTIPPVVALVMSRFGF---FAHLLPRCALVNSYLW  773
Db   800 F--LLLVLALPQQAYALDAAEQGELGL---VILVIISIFTLTPAYKILLSRSVWMLSYML 854

Qy   774 ------QRW------ENWFW-NVTLRPERFFLV--LVCFPGATYDALVTFCVCH       812
Db   855 VLEAEAQIQQWVPPLEARGGRDGIIWLAVILHPRLVFEVTKWLLAILGSAY---------  904

Qy   813 VALLCLTSSAASFFGTDSRVRAHRM----LVRLGKCHAWYSHYVLKFFLLVFGENGVFF  867
Db   905 --LLKASLLRIPYF-----VRAHALLRVCTLVR----HLAGARYIQMLLITIGRWTGYI  953

Qy   868 YKHLHGDVLPNDFASKLPLQ----------EP--FFPFEGKARVYRNEGRRLACGD     911
Db   954 YDHLS--------------PLSTWAAQGLRDLAVAVEPVVFSPMERKVIVWGAE--TVACGD 999

Qy   912 TVDGLPVVARLGDLVFAGLAMPPD------GWAITAPFTLQCLSERGTLSAMAVVMTGIDP 966
Db  1000 ILHGLPVSARLGREVLLG-----PADSYTSKGWKLLAPITAYTQQTRGLLGAIVVSLTGRDK 1056

Qy   967 RTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAANDQDIYQ 1026
Db  1057 NEQAGQVQVLSSVTQSFLGTSISGVLWTVYHGAGNKTLAGPKGPATQMYTSAEGDLVGWP  1116

Qy  1027 PPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCS 1086
Db  1117 SPPGTKSLDPCTCGAVDLYLVTRNADVIPVRRKDDRRGALLSPRPLSTLKGSSGGPVLCP 1176

Qy  1087 SGHVIGMFTAARNSGG---SVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILI 1143
Db  1177 RGHAVGLFRAAVCSRGVAKSIDFIPVESLDIARRTPSFSDNS---TPPAVPQSYQVGYLH 1233

Qy  1144 APTGSGKSTKLPLSMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYNGKCTNTGA  1203
Db  1234 APTGSGKSTKVPAAYASQGYKVLVLNPSVAATLGFGAYMSKAHGINPNIRTGVRTVTTGD 1293

Qy  1204 SLTYSTYGMYLT-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATA 1261
Db  1294 SITYSTYGKFLADGGCSAGAYDIIICDECHSVDATTILGIGTVLDQAETAGVRLVVLATA 1353

Qy  1262 TPPGVIPTPHANITEIQLTDEGTIPPHGKKIKEENLKKGRHLIFEATKKHCDELANELAR 1321
Db  1354 TPPGVTTPHSNIEEVALGHEGEIPFYGKAIPLAFIKGGRHLIFCHSKKKKCDELAAALRG 1413

Qy  1322 KGITAVSYYRGCDISKIP-EGDCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT  1380
Db  1414 MGVNAVAYYRGLDVSVIPTGQDVVVVATDALMTGYTGDFDSVIDCNVAVTQIVDFSLDPT 1473

Qy  1381 FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYG 1440
Db  1474 FTITTQTVPQDAVSRSQRRGRTGRGRLGTYRYVSSGERPSGMFDSVVLCECYDAGAAWYE 1533

Qy  1441 LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAA 1499
Db  1534 LTPAETTVRLRAYFNTPGLPVCQDHLEFWEAVFTGLTHIDAHFLSQTKQGGENFAYLTAY 1593

Qy  1500 QLQLCHQYGYAAPNDAPRWQGARLGKK-----PCGVLWRLGDADACPGPEPSEVTRY---- 1551
Db  1594 QATVCARAKAPPPSWDVTWKCLTRLKPTLTGPTPLLYRL-GAVTNEVTLTHPVTKYIATC 1652

Qy  1552 QMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCW---SITSVPTGATVAPVVDE 1608
Db  1653 MQADLEIMTS-SWVLAGGVLAAVAAYCLATGCISIIGRIHLNNQVVVTPDKEILYEAFDE 1711

Qy  1609 EEIVEECASFIPL-EAMVAAIDKLKSTI----TTTSPFTLETALE---KLNTFLGPH  1657
Db  1712 ---MEECASKAALIEEGQRIAEMLKSKIQGLLQQATRQAQDIQPAIQSSWPKLEQFWAKH 1768
```

```
Qy   2608 RVFASWMKVMGAPQDCVPQPKYSLEELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCS 2667
          |||  |        ::|:|:|| ::|||||||:  |: |:||||| |:: :
Db   2779 RAFTEAMTRYSAPPGDLPRPEYDLELITSCSSNVSVALDSRGRRRYYLTRDPTTPITRAA 2838

Qy   2668 AEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPV 2727
          : :||:  || |:|||  ||| |:  : |  :| :||  :| |:|  :|:|
Db   2839 WETVRHSPVNSWLGNIIQYAPTIWVRMVIMTHFFAILLAQDTLNQNLNFEMYGAVYSVNP 2898

Qy   2728 EDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGG 2787
          ||||:|     :|||:|| |: |: :|   ||||||        |||||    ::||||
Db   2899 LDLPAIIERLHGLEAFSLHTYSPHELSRVAATLRKLGAPPLRAWKSRARAVRASLIAQGG 2958

Qy   2788 AHAKLARFLL-WHATSR---PLPDLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKF 2842
          ||| |    :: :   |  :|:|||  || : ||    |   |:|:||  :
Db   2959 RAATCGRYLFNWAVRTKLKLTPLPEASRLDLSGWFT-----VGAGGGDIFHSVSHARPRL 3013

Qy   2843 LVKYLAVIVFALGLIAVGLAI 2863
          :      | |:|||:|
Db   3014 L------LLCLLLSVGVGI 3027
```

Search completed: October 27, 2005, 15:46:51
Job time : 354 secs

```
                    GenCore version 5.1.6
             Copyright (c) 1993 - 2005  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         October 27, 2005, 15:32:46 ; Search time 64 Seconds
                              (without alignments)
                              3340.546 Million cell updates/sec

Title:          US-10-009-002-2
Perfect score:  15376
Sequence:       1 MPVISTQTSPVPAPRTRKNK.........KYLAVIVFALGLIAVGLAIS 2864

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters:          513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      Issued Patents AA:*
                1: /cgn2_6/ptcdata/1/iaa/5A_COMB.pep:*
                2: /cgn2_6/ptcdata/1/iaa/5B_COMB.pep:*
                3: /cgn2_6/ptcdata/1/iaa/6A_COMB.pep:*
                4: /cgn2_6/ptcdata/1/iaa/6B_COMB.pep:*
                5: /cgn2_6/ptcdata/1/iaa/PCTUS_COMB.pep.*
                6: /cgn2_6/ptcdata/1/iaa/backfiles1.pep:*

         Pred. No. is the number of results predicted by chance to have a
         score greater than or equal to the score of the result being printed,
         and is derived by analysis of the total score distribution.
```

SUMMARIES

```
                 %
Result           Query
No.    Score  Match Length DB  ID                        Description
------ -----  ----- ------ --  -----------------------   ------------------
  1    15317  99.6   2864   4  US-08-469-260A-394        Sequence 394, App
  2    15317  99.6   2864   4  US-08-488-446-394         Sequence 394, App
  3    15317  99.6   2864   4  US-08-467-344A-394        Sequence 394, App
  4    15317  99.6   2864   4  US-08-424-550B-394        Sequence 394, App
  5     7380  48.0   1422   4  US-08-469-260A-83         Sequence 83, Appl
  6     7380  48.0   1422   4  US-08-488-446-83          Sequence 83, Appl
  7     7380  48.0   1422   4  US-08-467-344A-83         Sequence 83, Appl
  8     7380  48.0   1422   4  US-08-424-550B-83         Sequence 83, Appl
  9   3274.5  21.3   3033   1  US-07-925-695-8           Sequence 8, Appli
 10    3269  21.3   3033   1  US-07-925-695-5           Sequence 5, Appli
 11    3262  21.2   3010   4  US-09-539-601-3           Sequence 3, Appli
 12    3261  21.2   3010   4  US-09-539-601-33          Sequence 33, Appl
 13    3255  21.1   3010   4  US-09-539-601-27          Sequence 27, Appl
 14    3252  21.1   3010   4  US-09-539-601-21          Sequence 21, Appl
 15    3247  21.1   3010   3  US-09-014-416-3           Sequence 3, Appli
 16   3230.5  21.0   3033   1  US-07-925-695-9           Sequence 9, Appli
 17    3218  21.0   3010   1  US-08-324-977-2           Sequence 2, Appli
 18    3218  20.9   3010   1  US-08-324-977-14          Sequence 14, Appl
 19    3218  20.9   3010   2  US-08-384-616-2           Sequence 2, Appli
 20    3218  20.9   3010   2  US-08-384-616-14          Sequence 14, Appl
 21    3218  20.9   3010   2  US-08-904-686A-2          Sequence 2, Appli
 22    3218  20.9   3010   3  US-08-904-686A-14         Sequence 14, Appl
 23    3218  20.9   3010   3  US-09-315-850-2           Sequence 2, Appli
 24    3218  20.9   3010   3  US-09-315-850-14          Sequence 14, Appl
 25   3215.5  20.9   3011   3  US-08-850-328-1          Sequence 1, Appli
 26   3214.5  20.9   3011   3  US-09-014-416-1           Sequence 1, Appli
 27   3214.5  20.9   3011   4  US-09-952-572-9           Sequence 9, Appli
 28          20.9   3209.5  3011  1  US-08-188-281B-1    Sequence 1, Appli
 29          20.9   3205.5  3011  1  US-08-453-552-1     Sequence 1, Appli
 30          20.9   3205.5  3011  2  US-08-710-637-1     Sequence 1, Appli
 31          20.9   3209.5  3011  5  PCT-US93-00907-1    Sequence 1, Appli
 32          20.9   3205.5  3011  5  PCT-US94-07280-1    Sequence 1, Appli
 33          20.9   3205.5  3011  5  PCT-US95-01087-1    Sequence 2, Appli
 34          20.9   3208.5  3012  3  US-08-811-566-2     Sequence 2, Appli
 35          20.8   3208.5  3012  3  US-09-034-756-2     Sequence 36, Appl
 36          20.8   3204.5  3011  1  US-08-440-103-36    Sequence 36, Appl
 37          20.8   3204.5  3011  1  US-08-440-542-36    Sequence 10, Appl
 38          20.8   3204.5  3011  1  US-07-910-760-10    Sequence 10, Appl
 39          20.8   3204.5  3011  1  US-08-440-519-10    Sequence 36, Appl
 40          20.8   3204.5  3011  1  US-08-231-368-36    Sequence 36, Appl
 41          20.8   3204.5  3011  1  US-08-440-210-36    Sequence 2, Appli
 42          20.8   3204.5  3011  3  US-09-388-874-2     Sequence 36, Appl
 43          20.8   3204.5  3011  3  US-09-046-604-36    Sequence 10, Appl
 44          20.8   3204.5  3011  3  US-08-440-549-10    Sequence 2, Appli
 45          20.8   3204.5  3011  4  US-09-916-359-2     Sequence 2, Appli
```

ALIGNMENTS

```
RESULT 1
US-08-469-260A-394
; Sequence 394, Application US/08469260A
; Patent No. 6451578
; GENERAL INFORMATION:
;  APPLICANT:  JOHN N.  SIMONS
;  APPLICANT:  TAMI J.  PILOT-MATIAS
;  APPLICANT:  GEORGE J.  DAWSON
;  APPLICANT:  GEORGE G.  SCHLAUDER
;  APPLICANT:  SURESH M.  DESAI
;  APPLICANT:  THOMAS P.  LEARY
;  APPLICANT:  ANTHONY SCOTT MUERHOFF
;  APPLICANT:  JAMES C.  ERKER
;  APPLICANT:  SHERI L.  BUIJK
;  APPLICANT:  ISA K.  MUSHAHWAR
;  TITLE OF INVENTION:  NON-A, NON-B. NON-C. NON-D, NON-E HEPATITIS
;  TITLE OF INVENTION:  REAGENTS AND METHODS FOR THEIR USE
;  NUMBER OF SEQUENCES:  716
;  CORRESPONDENCE ADDRESS:
;    ADDRESSEE:  ABBOTT LABORATORIES  D377/AP6D
;    STREET:  100 ABBOTT PARK ROAD
;    CITY:  ABBOTT PARK
;    STATE:  IL
;    COUNTRY:  USA
;    ZIP:  60064-3500
;  COMPUTER READABLE FORM:
;    MEDIUM TYPE:  Floppy disk
;    COMPUTER:  IBM PC compatible
;    OPERATING SYSTEM:  PC-DOS/MS-DOS
;    SOFTWARE:  PatentIn Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;    APPLICATION NUMBER:  US/08/469,260A
;    FILING DATE:
;    CLASSIFICATION:
;  PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  US/08/424,550
;    FILING DATE:
;  ATTORNEY/AGENT INFORMATION:
;    NAME:  POREMBSKI, PRISCILLA E.
;    REGISTRATION NUMBER:  33,207
;    REFERENCE/DOCKET NUMBER:  5527.PC.01
;  TELECOMMUNICATION INFORMATION:
;    TELEPHONE:  708-937-6365
;    TELEFAX:  708-938-2623
; INFORMATION FOR SEQ ID NO:  394:
;  SEQUENCE CHARACTERISTICS:
;    LENGTH:  2864 amino acids
;    TYPE:  amino acid
;    TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
```

```
US-08-469-260A-394

Query Match        99.6%;   Score 15317;   DB 4;   Length 2864;
Best Local Similarity 99.6%;   Pred. No. 0;
Matches 2852;  Conservative 4;  Mismatches 4;  Indels 8;  Gaps 0;

Qy    1 MPVISTQTSPVPAPRTRNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MPVISTQTSPVPAPRTRNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD 60

Qy   61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPV 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPV 120

Qy  121 CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180

Qy  181 CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240

Qy  241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK 300

Qy  301 LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI 360

Qy  361 AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF 420
          |||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db  361 AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPVTLEYNNSISWYPYTIPGARGCMVKF 420

Qy  421 KNNTWGGCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAIILQYPG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 KNNTWGGCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAIILQYPG 480

Qy  481 SKEMFKPHNWMSGHLYEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 SKEMFKPHNWMSGHLYEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ 540

Qy  541 GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY 600

Qy  601 LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL 660

Qy  661 LGFVPMAAGLPLTFVFVAAAAAQPDYDWMWVRLLVAGRNRGHRIALLVGPWPLVAL 720
          |||||||||||||||||||||||||||||||||||||| :|| :|||||||||||
Db  661 LGFVPMAAGLPLTFVFVAAAAAQPDYDWMWVRLLVAGRDRGPRIALLVGPWPLVAL 720

Qy  721 LTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRGFFAHLLPRCALVNSYLMQRWENWF 780
          ||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 LTLLHLATPASAFDTEIIGGLTIPPVVALVVMSRGFFAHLLPRCALVNSYLMQRWENWF 780

Qy  781 WNVTLRPERFFLVLVCFPGATYDALVTFCVCHAVLCLTSSAASFFGTDSRVRAHRMLVR 840
          |||||||||||:||||||||||| :|||||||||||||||||||||||||||||||||
Db  781 WNVTLRPERFFLLVLVCFPGATYDTLVTFCVCHAVLCLTSSAASFFGTDSRVRAHRMLVR 840

Qy  841 LGKCHAWYSHYVLKFFLLVFGENGVFFYKGHLHGDVLYTAHHGSKGRRLAHPTGSIHPTE 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 LGKCHAWYSHYVLKFFLLVFGENGVFFYKGHLHGDVLYTAHHGSKGRRLAHPTGSIHPTE 900

Qy  901 RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPPTLQCLSERGTLSAMAVV 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPPTLQCLSERGTLSAMAVV 960

Qy  961 MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020


Qy 1021 DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG 1080

Qy 1081 APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ 1140

Qy 1141 ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200

Qy 1201 TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT 1260

Qy 1261 ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA 1320

Qy 1321 RKGITAVSYRGCDISKIPEGDCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 RKGITAVSYRGCDISKIPEGDCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT 1380

Qy 1381 FTMGVRVCGVSAIVKGQRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 FTMGVRVCGVSAIVKGQRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYG 1440

Qy 1441 LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ 1500

Qy 1501 LQLCHQYGYAAPNDAPRWQGARLGKKPCCVLWRLDGADDACPGPEPSEVTRYQMCFTEVNT 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 LQLCHQYGYAAPNDAPRWQGARLGKKPCCVLWRLDGADDACPGPEPSEVTRYQMCFTEVNT 1560

Qy 1561 SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEECASFIP 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEECASFIP 1620

Qy 1621 LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA 1680

Qy 1681 SCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 SCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF 1740

Qy 1741 VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT 1800

Qy 1801 AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE 1860

Qy 1861 DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ 1920

Qy 1921 ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVNYG 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db 1921 ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVVNYG 1980

Qy 2041 CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE 2100
          ||||||||||||||||||||||||||||||||||||||| :|||||||||||||||||
Db 2041 CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSINNTPSDEAAVSALVFKQE 2100
```

```
                    STREET: 100 ABBOTT PARK ROAD
                    CITY: ABBOTT PARK
                    STATE: IL
                    COUNTRY: USA
                    ZIP: 60064-3500
          COMPUTER READABLE FORM:
                    MEDIUM TYPE: Floppy disk
                    COMPUTER: IBM PC compatible
                    OPERATING SYSTEM: PC-DOS/MS-DOS
                    SOFTWARE: PatentIn Release #1.0, Version #1.25
          CURRENT APPLICATION DATA:
                    APPLICATION NUMBER: US/08/488,446
                    FILING DATE:
                    CLASSIFICATION:
          PRIOR APPLICATION DATA:
                    APPLICATION NUMBER: US/08/424,550
                    FILING DATE:
          ATTORNEY/AGENT INFORMATION:
                    NAME: POREMBSKI, PRISCILLA E.
                    REGISTRATION NUMBER: 33,207
                    REFERENCE/DOCKET NUMBER: 5527.PC.01
          TELECOMMUNICATION INFORMATION:
                    TELEPHONE: 708-937-6365
                    TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 394:
          SEQUENCE CHARACTERISTICS:
                    LENGTH: 2864 amino acids
                    TYPE: amino acid
                    TOPOLOGY: linear
          MOLECULE TYPE: protein
US-08-488-446-394

Query Match             99.6%;  Score 15317;  DB 4;  Length 2864;
Best Local Similarity   99.6%;  Pred. No. 0;
Matches 2852;  Conservative 4;  Mismatches 8;  Indels 0;  Gaps 0;

Qy    1 MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD 60

Qy   61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPV 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPV 120

Qy  121 CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180

Qy  181 CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240

Qy  241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK 300

Qy  301 LASQVPYAIATMFSSVHYLAVGALIIYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LASQVPYAIATMFSSVHYLAVGALIIYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI 360

Qy  361 ABFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF 420
          |||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db  361 ABFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPVTLEYNNSISWYPYTIPGARGCMVKF 420

Qy  421 KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTAWHNGSALKLAILQYPG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTAWHNGSALKLAILQYPG 480

Qy  481 SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ 540

Qy  541 GFYSDVKDLATGLIITKDKAWKNYQVLLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY 600
```

```
Qy 2101 LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL 2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101 LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL 2160

Qy 2161 QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK 2220
          |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db 2161 QRSDPLEGPSNLPSSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK 2220

Qy 2221 EVSEWSDEWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCCMSYTW 2280
          |||||||:|:||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221 EVSEWSDGSWSTTTTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCCMSYTW 2280

Qy 2281 TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ 2340
          |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db 2281 TDVISFKTASKVLSATRAITSGFLKQRSLVVTEPRDAELRKQKVTINRQPLFPPSYHKQ 2340

Qy 2341 VRLAKEKASKVGVMMDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA 2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2341 VRLAKEKASKVGVMMDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA 2400

Qy 2401 GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV 2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401 GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV 2460

Qy 2461 MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH 2520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2461 MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH 2520

Qy 2521 RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAABQAGMK 2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2521 RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAABQAGMK 2580

Qy 2581 NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN 2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2581 NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN 2640

Qy 2641 VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF 2700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2641 VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF 2700

Qy 2701 MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL 2760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2701 MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL 2760

Qy 2761 TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN 2820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761 TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN 2820

Qy 2821 YCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS 2864
          |||||||||||:||||||||||||||||||||||||||||||||
Db 2821 YCDVYSPEGDVFVTPQRRLQKFLVKYLAVIVFALGLIAVGLAIS 2864
```

Right column (residues 1621–2760):

```
Db   1621  LEAMVAAIDKLKSTITTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFFA  1680
Qy   1681  SCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAGTALGTWTSVGF   1740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681  SCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAGTALGTWTSVGF   1740
Qy   1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVGVLSACAMFALTT   1800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVGVLSACAMFALTT   1800
Qy   1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE  1860
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE  1860
Qy   1861  DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ  1920
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861  DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ  1920
Qy   1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDTDWTSLVVNYG   1980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDTDWTSLVVNYG   1980
Qy   1981  VRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVADVGVQCYLGEPKTPWTTSAC    2040
           ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981  VRDYCKYEKLGDHIFVTAVSSPNVCFTQVPPTLRAAVADVGVQCYLGEPKTPWTTSAC    2040
Qy   2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE  2100
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSINNTPSDEAAVSALVFKQE  2100
Qy   2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL   2160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL   2160
Qy   2161  QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK  2220
           ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   2161  QRSDPLEGPSNLPSSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK  2220
Qy   2221  EVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280
           |||||||| ||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   2221  EVSEWSDGSWSTTTTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280
Qy   2281  TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281  TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340
Qy   2341  VRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341  VRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400
Qy   2401  GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRRLISYPHLEMRCVEKMYYGQVAPDVVKAV 2460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401  GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRRLISYPHLEMRCVEKMYYGQVAPDVVKAV 2460
Qy   2461  MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461  MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520
Qy   2521  VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF  2580
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521  VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF  2580
Qy   2581  NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN  2640
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581  NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN  2640
Qy   2641  RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK  2700
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641  RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK  2700
Qy   2701  MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL  2760
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701  MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL  2760
```

Left column (residues 541–1680):

```
Db         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541  GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYILLAY  600
Db    601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYLRCRLRYAAL  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYLRCRLRYAAL  660
Db    661  LGFVPMAAGLPLTFFVAAAAAQPDYDMWVRLLVAGRDRGPRIALLVGPWPLVAL        720
           |||||||||||||||||||||||||||||||||||| :|||| ||||||||||||
Qy    661  LGFVPMAAGLPLTFFVAAAAAQPDYDMWVRLLVAGRNRGHRIALLVGPWPLVAL        720
Db    721  LTLLHLATPASAFDTEIIGGLTIPPVVALVVNSRFGFFAHLLPRCALVNSYLWQRWENWF  780
           ||||| |||||||||||||||||||||||||| |||||||||||||||||||||||||||
Qy    721  LTLLHVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQRWENWF   780
Db    781  WNVTLRPERPFLLVLVCFPGATYDTLVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR 840
           ||||||||||| |||||| ||| ||| |||||||||||||||||||||||||||||||||
Qy    781  WNVTLRPERPFLVLVCFPGATYDALVTFCVCHVALLCLTSSAASFGTDSRVRAHRMLVR   840
Db    841  LGKCHAWYSHYVLKFPLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPFEGKARVY  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    841  LGKCHAWYSHYVLKFPLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPFEGKARVY  900
Db    901  RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPFTLQCLSERGTLSAMAVV  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    901  RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPFTLQCLSERGTLSAMAVV  960
Db    961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN  1020
Db   1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG  1080
Db   1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ  1140
Db   1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN  1200
Db   1201  TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1201  TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT  1260
Db   1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA  1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA  1320
Db   1321  RKGITAVSYYRGCDISKIPEGDCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT   1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1321  RKGITAVSYYRGCDISKIPEGDCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT   1380
Db   1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEARFDAAKAWYG 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEARFDAAKAWYG 1440
Db   1441  LSSTEAQTILDYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ   1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1441  LSSTEAQTILDYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ   1500
Db   1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT  1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT  1560
Db   1561  SGTAALAVGVGVAMAYLAIDFFGATCVRRCWSITSVPTGATVAPVVDEEIVEECASFIP   1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1561  SGTAALAVGVGVAMAYLAIDFFGATCVRRCWSITSVPTGATVAPVVDEEIVEECASFIP   1620
Db   1621  LEAMVAAIDKLKSTITTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFFA  1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1621  LEAMVAAIDKLKSTITTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFFA  1680
```

Qy   61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPV 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPV 120

Qy 121 CQIVRLLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 CQIVRLLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180

Qy 181 CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240

Qy 241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK 300

Qy 301 LASQVPYAIATMPSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 LASQVPYAIATMPSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI 360

Qy 361 AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF 420
         |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db 361 AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPVTLEYNNSISWYPYTIPGARGCMVKF 420

Qy 421 KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG 480

Qy 481 SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ 540

Qy 541 GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY 600

Qy 601 LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFCCYLRCRLRYAAL 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFCCYLRCRLRYAAL 660

Qy 661 LGFVPMAAGLPLTFFVAAAAQPDYDWMWVRLLVAGLVLWAGNRGHRIALLVGPWPLVAL 720
         |||||||||||||||||||||||||||||||||||||||| |:|| |||||||||||||
Db 661 LGFVPMAAGLPLTFFVAAAAQPDYDWMWVRLLVAGLVLWAGRDRGPRIALLVGPWPLVAL 720

Qy 721 LTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFPAHLLPRCALVNSYLWQRWENWF 780
         ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 LTLLHLATPASAFDTEIIGGLTIPPVVALVVMSRFGFPAHLLPRCALVNSYLWQRWENWF 780

Qy 781 WNVTLRPERFLLVLCFPGATYDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR 840
         ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db 781 WNVTLRPERFLLVLCFPGATYDTLVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR 840

Qy 841 LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQSPFPPEGKARVY 900
         |||||||||||||||||||||||||||||||||||||||||||||| ||| |||||||
Db 841 LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPEGKARVY 900

Qy 901 RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPPTLQCLSERGTLSAMAVV 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPPTLQCLSERGTLSAMAVV 960

Qy 961 MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020

Qy 1021 DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG 1080

Qy 1081 APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ 1140

Qy 1141 ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200

Qy 2761 TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761 TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN 2820

Qy 2821 YCDVVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS 2864
         |||||||||||||: |||||||||||||||||||||||||||||
Db 2821 YCDVVYSPEGDVFTPQRRLQKFLVKYLAVIVFALGLIAVGLAIS 2864

RESULT 3
US-08-467-344A-394
; Sequence 394, Application US/08467344A
; Patent No. 6586568
GENERAL INFORMATION:
     APPLICANT: JOHN N. SIMONS
                TAMI J. PILOT-MATIAS
                GEORGE J. DAWSON
                GEORGE G. SCHLAUDER
                SURESH M. DESAI
                THOMAS P. LEARY
                ANTHONY SCOTT MUERHOFF
                JAMES C. ERKER
                SHERI L. BUIJK
                ISA K. MUSHAHWAR
     TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
                REAGENTS AND METHODS FOR THEIR USE
     NUMBER OF SEQUENCES: 716
     CORRESPONDENCE ADDRESS:
          ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
          STREET: 100 ABBOTT PARK ROAD
          CITY: ABBOTT PARK
          STATE: IL
          COUNTRY: USA
          ZIP: 60064-3500
     COMPUTER READABLE FORM:
          MEDIUM TYPE: Floppy disk
          COMPUTER: IBM PC compatible
          OPERATING SYSTEM: PC-DOS/MS-DOS
          SOFTWARE: PatentIn Release #1.0, Version #1.25
     CURRENT APPLICATION DATA:
          APPLICATION NUMBER: US/08/467,344A
          FILING DATE: 07-Jun-1995
          CLASSIFICATION: <Unknown>
     PRIOR APPLICATION DATA:
          APPLICATION NUMBER: 08/424,550
          FILING DATE: <Unknown>
     ATTORNEY/AGENT INFORMATION:
          NAME: POREMBSKI, PRISCILLA E.
          REGISTRATION NUMBER: 33,207
          REFERENCE/DOCKET NUMBER: 5527.PC.01
     TELECOMMUNICATION INFORMATION:
          TELEPHONE: 708-937-6365
          TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 394:
     SEQUENCE CHARACTERISTICS:
          LENGTH: 2864 amino acids
          TYPE: amino acid
          TOPOLOGY: linear
          MOLECULE TYPE: protein
          SEQUENCE DESCRIPTION: SEQ ID NO: 394:
US-08-467-344A-394

Query Match   99.6%; Score 15317; DB 4; Length 2864;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2852; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy   1 MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD 60

Qy 61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPV 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Qy  1201  TGASLYTYSTYGYMLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  TGASLYTYSTYGYMLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT  1260

Qy  1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA  1320

Qy  1321  RKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  RKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT  1380

Qy  1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYG  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYG  1440

Qy  1441  LSSTEAQTILDYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ  1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  LSSTEAQTILDYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ  1500

Qy  1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT  1560

Qy  1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEECASFIP  1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEECASFIP  1620

Qy  1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA  1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA  1680

Qy  1681  SCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  SCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF  1740

Qy  1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT  1800

Qy  1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVSACIRWLHTPTE  1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVSACIRWLHTPTE  1860

Qy  1861  DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ  1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ  1920

Qy  1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVNYG  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVNYG  1980

Qy  1981  VRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDGVQVQCYLGEPKTPWTTSAC  2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  VRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDGVQVQCYLGEPKTPWTTSAC  2040

Qy  2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCSNTNTPSDEAAVSALVFKQE  2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCSNTNTPSDEAAVSALVFKQE  2100

Qy  2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL  2160
          ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db  2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL  2160

Qy  2161  QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK  2220
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db  2161  QRSDPLEGPSNLPSSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK  2220

Qy  2221  EVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280
          ||||||||:||||  |||||||||||||||||||||||||||||||||||||||||||||
Db  2221  EVSEWSDGSWSTTTTTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280


Qy  2281  TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340

Qy  2341  VRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  VRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400

Qy  2401  GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV  2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV  2460

Qy  2461  MGDAYGFVDPRITRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  MGDAYGFVDPRITRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520

Qy  2521  RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK  2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK  2580

Qy  2581  NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN  2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN  2640

Qy  2641  VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF  2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF  2700

Qy  2701  MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL  2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL  2760

Qy  2761  TDMTMPPLRAWRKKARAVLASAKRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN  2820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  TDMTMPPLRAWRKKARAVLASAKRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN  2820

Qy  2821  YCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS  2864
          ||||||||||||:|||||||||||||||||||||||||||||||
Db  2821  YCDVYSPEGDVFTPQRRLQKFLVKYLAVIVFALGLIAVGLAIS  2864
```

```
;    FILING DATE:
;    CLASSIFICATION:  435435
;    ATTORNEY/AGENT INFORMATION:
;    NAME:  POREMBSKI, PRISCILLA E.
;    REGISTRATION NUMBER:  33,207
;    REFERENCE/DOCKET NUMBER:  5527.PC.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708-937-6365
;      TELEFAX:  708-938-2623
;    INFORMATION FOR SEQ ID NO:  394:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2864 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-424-550B-394

Query Match         99.6%;  Score 15317;  DB 4;  Length 2864;
Best Local Similarity  99.6%;  Pred. No. 0;
Matches 2852;  Conservative  4;  Mismatches  4;  Indels  8;  Gaps  0;
```

```
Qy       1  MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRARKKVQRDARPRNYKIAGIHD   60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1  MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRARKKVQRDARPRNYKIAGIHD   60

Qy      61  GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPV  120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61  GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPV  120

Qy     121  CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSSGARVTDPDTNTTILTNCCQRNQVIY  180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121  CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSSGARVTDPDTNTTILTNCCQRNQVIY  180

Qy     181  CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG  240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181  CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG  240

Qy     241  ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK  300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241  ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK  300

Qy     301  LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI  360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301  LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI  360

Qy     361  AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF  420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361  AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPVTLEYNNSISWYPYTIPGARGCMVKF  420

Qy     421  KNNTWGCCRIRRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWINGSALKLAIIQYPG  480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421  KNNTWGCCRIRRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAIIQYPG  480

Qy     481  SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ  540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481  SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ  540

Qy     541  GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY  600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541  GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY  600

Qy     601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYLRCRLRYAAL  660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYLRCRLRYAAL  660

Qy     661  LGFVPMAAGLPLTFFVAAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALLVGPWPLVAL  720
            |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||:||
Db     661  LGFVPMAAGLPLTFFVAAAAAQPDYDWWVRLLVAGLVLWAGRDRGPRIALLVGPWPLVAL  720

Qy     721  LTLLHLVTPASAFDTEIGGLITIPPVVALVVMSRFGFFAHLLPRCALVNSYLMQRWENWF  780
            ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721  LTLLHLATPASAFDTEIGGLITIPPVVALVVMSRFGFFAHLLPRCALVNSYLMQRWENWF  780
```

```
Qy     781  WNVTLRPERFPLVLVCFPGATYDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR  840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781  WNVTLRPERFPLVLVCFPGATYDTLVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR  840

Qy     841  LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPEGKARVY  900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841  LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPEGKARVY  900

Qy     901  RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPPTLQCLSERGTLSAMAVV  960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901  RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPPTLQCLSERGTLSAMAVV  960

Qy     961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020

Qy    1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG 1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG 1080

Qy    1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ 1140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ 1140

Qy    1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200

Qy    1201  TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT 1260
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201  TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT 1260

Qy    1261  ATPPGVIPTPHANITEIQLTDEGTIPHGKKIKEENLKKGRHLIFEATKKHCDELANELA 1320
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261  ATPPGVIPTPHANITEIQLTDEGTIPHGKKIKEENLKKGRHLIFEATKKHCDELANELA 1320

Qy    1321  RKGITAVSYYRGCDISKIPEGDCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT 1380
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321  RKGITAVSYYRGCDISKIPEGDCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT 1380

Qy    1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCCTPSGMVPECNIVEAFDAAKAWYG 1440
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCCTPSGMVPECNIVEAFDAAKAWYG 1440

Qy    1441  LSSTEAQTILDYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ 1500
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441  LSSTEAQTILDYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ 1500

Qy    1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT 1560
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT 1560

Qy    1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEBECASFIP 1620
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEBECASFIP 1620

Qy    1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA 1680
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA 1680

Qy    1681  SCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAGTALGTWTSVGF 1740
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681  SCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAGTALGTWTSVGF 1740

Qy    1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT 1800
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT 1800

Qy    1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE 1860
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE 1860

Qy    1861  DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGVKGPWIGSGMLQ 1920
```

APPLICANT: JOHN N. SIMONS
APPLICANT: TAMI J. PILOT-MATIAS
APPLICANT: GEORGE J. DAWSON
APPLICANT: GEORGE G. SCHLAUDER
APPLICANT: SURESH M. DESAI
APPLICANT: THOMAS P. LEARY
APPLICANT: ANTHONY SCOTT MUERHOFF
APPLICANT: JAMES C. ERKER
APPLICANT: SHERI L. BUIJK
APPLICANT: ISA K. MUSHAHWAR
TITLE OF INVENTION: NON-A, NON-B. NON-C, NON-D, NON-E HEPATITIS
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
NUMBER OF SEQUENCES: 716
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES   D377/AP6D
STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,260A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/424,550
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: POREMBSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5527.PC.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 1422 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-469-260A-83

Query Match   48.0%;   Score 7380;   DB 4;   Length 1422;
Best Local Similarity 97.3%;   Pred. No. 0;
Matches 1392; Conservative 27; Mismatches 2; Indels 10; Gaps 2;

```
Qy   1005  AHPTGSIHPITVDAANDDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYW 1064
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  AHPTGSIHPITVDAANDDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYW   60

Qy   1065  CVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAH 1124
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVGQIRVRPLVCAGYHPQYTAH  120

Qy   1125  ATLDTKPTVPNEYSVQILLAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHA 1184
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  ATLDTKPTVPNEYSVQILLAPTGSGKSTKLPLSYMQXKKEVLVLNPSVATTASMPKYMHA  180

Qy   1185  TYGVNPNCYFNGKCTNTGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKV 1244
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  TYGVNPNCYFNGKCTNTGASLTYSTYGMYLTGCSRNYDVIICDECHATDRTTVLGIGKV  240

Qy   1245  LTEAPSKNVRLVVLATATPPGVITPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLI 1304
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  LTEAPSKNVRLVVLATATPPGVITPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLI  300

Qy   1305  FEATKKHCDELANELARKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYD 1364
```

```
Db   1861  DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDTDWTSLVVNYG 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDTDWTSLVVNYG 1980

Qy   1981  VRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVADGVQVQCYLGEPKTPWTTSAC 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981  VRDYCKYEKLGDHIFVTAVSSPNVCFTQVPPTLRAAVADGVQVQCYLGEPKTPWTTSAC 2040

Qy   2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE 2100

Qy   2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL 2160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL 2160

Qy   2161  QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGPDDLPSYPPKK 2220
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161  QRSDPLEGPSNLPSSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGPDDLPSYPPKK 2220

Qy   2221  EVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221  EVSEWSDGSWSTTTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW 2280

Qy   2281  TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ 2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281  TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ 2340

Qy   2341  VRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA 2400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341  VRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA 2400

Qy   2401  GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV 2460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401  GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV 2460

Qy   2461  MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH 2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461  MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH 2520

Qy   2521  RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK 2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521  RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK 2580

Qy   2581  NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN 2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581  NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN 2640

Qy   2641  VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF 2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641  VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF 2700

Qy   2701  MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL 2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701  MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL 2760

Qy   2761  TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTFFN 2820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761  TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN 2820

Qy   2821  YCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS 2864
           ||||||||||||||||||||||||||||||||||||||||||||
Db   2821  YCDVYSPEGDVFTPQRRLQKFJVKYLAVIVFALGLIAVGLAIS 2864
```

RESULT 5
US-08-469-260A-83
; Sequence 83, Application US/08469260A
; Patent No. 6451578
; GENERAL INFORMATION:

```
Db   1379 ------VLDLQKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLI 1422
```

RESULT 6
US-08-488-446-83
; Sequence 83, Application US/08488446
; Patent No. 6558898
GENERAL INFORMATION:
   APPLICANT: JOHN N. SIMONS
   APPLICANT: TAMI J. PILOT-MATIAS
   APPLICANT: GEORGE J. DAWSON
   APPLICANT: GEORGE G. SCHLAUDER
   APPLICANT: SURESH M. DESAI
   APPLICANT: THOMAS P. LEARY
   APPLICANT: ANTHONY SCOTT MUERHOFF
   APPLICANT: JAMES C. ERKER
   APPLICANT: SHERI L. BUIJK
   APPLICANT: ISA K. MUSHAHWAR
   TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
   TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
   NUMBER OF SEQUENCES: 716
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: ABBOTT LABORATORIES   D377/AP6D
      STREET: 100 ABBOTT PARK ROAD
      CITY: ABBOTT PARK
      STATE: IL
      COUNTRY: USA
      ZIP: 60064-3500
   COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.25
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/488,446
      FILING DATE:
      CLASSIFICATION:
   PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US/08/424,550
      FILING DATE:
   ATTORNEY/AGENT INFORMATION:
      NAME: POREMBSKI, PRISCILLA E.
      REGISTRATION NUMBER: 33,207
      REFERENCE/DOCKET NUMBER: 5527.PC.01
   TELECOMMUNICATION INFORMATION:
      TELEPHONE: 708-937-6365
      TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 83:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 1422 amino acids
      TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
   MOLECULE TYPE: protein
US-08-488-446-83

Query Match   48.0%; Score 7380; DB 4; Length 1422;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1392; Conservative 2; Mismatches 27; Indels 10; Gaps 2;

Qy 1005 AHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCSCGETKGYLVTRLGGSLVEVNKSDDPYW 1064
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 AHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCSCGETKGYLVTRLGGSLVEVNKSDDPYW 60

Qy 1065 CVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAH 1124
        ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db   61 CVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVGQIRVRPLVCAGYHPQYTAH 120

Qy 1125 ATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHA 1184
        ||||||||||||||||||||||||||||||||||||:|:||||||||||||||||||||
Db 121 ATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQKKXEVLVLNPSVATTASMPKYMHA 180

Qy 1185 TYGVNPNCYFNGKCTNTGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKV 1244
```

```
Db   301 FEATKKHCDELANELARKGITAVSYYRGCDISKMPEGDCVVVATDALCTGYTGDFDSVYD 360
         |||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Qy 1365 CSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVP 1424
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVP 420

Qy 1425 ECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVN 1484
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 ECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVN 480

Qy 1485 TAKRTADNYVLLTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPE 1544
         |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db   481 TAKRTADNYVLLTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGCDACPGPE 540

Qy 1545 PSEVTRYQMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAP 1604
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 PSEVTRYQMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAP 600

Qy 1605 VVDEEIVEECASFIPLEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAI 1664
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 VVDEEIVEECASFIPLEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAI 660

Qy 1665 IEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMM 1724
         |||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db   661 IEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARXALAFMM 720

Qy 1725 AGAAGTALGTWTSVGFVFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAA 1784
         |||:||||||||||||||||||||||:||||||||||||||||:|||||||||||||||||
Db   721 AGAXGTALGTWTSVGFVFDMLGGYAGASSTACLTFKCLMGEWXTMDQLAGLVYSAFNPAA 780

Qy 1785 GVVGVLSACAMFALTTAGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTP 1844
         |||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db   781 GVVGVLSACAMFALTTAGPDHWPNRLLTMLARSNTVCXEYFIATRDIRRKILGILEASTP 840

Qy 1845 WSVISACIRWLHTPTEDDCGLIAWGLEIWQYVCNFVICFNVLKAGVQSMVNIPGCPFYS 1904
         ||:|||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db   841 WSXISACIRWLHTPTEDDCGLIAWGLXIWQYVCNFVICFNVLKAGVQSMVNIPGCPFYS 900

Qy 1905 CQKGYKGPWIGGSMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGS 1964
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CQKGYKGPWIGGSMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGS 960

Qy 1965 ARPDPTDWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDGVQV 2024
         |||||||||||||||||||||||||||:||||||||||||||||||||||||||||:|||
Db   961 ARPDPTDWTSLVVNYGVRDYCKYEKLGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDRVQV 1020

Qy 2025 QCYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNN 2084
         |:|||||||||||||||||||||||||||||||||||||:||||||||::||||||||:||
Db  1021 QXYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGGRMQLNLRDRLLEANDCNSINN 1080

Qy 2085 TPSDERAAVSALVFKQELRRTNQLLEAISAGVDTTKLPAPS-IEEVVRKRQFRARTGSLT 2143
         ||||||||||||||||||||||||||||||||||||||||| ||||||:||||||||||
Db  1081 TPSDERAAVSALVFKQELRRTNQLLEAISAGVDTTKLPAPSQIEEVVVRKRQFRARTGSLT 1140

Qy 2144 LPPPPRSVPGVSCPESLQRSDPLEGPSNLPPSSPPVLQLAMPMPLLGAGECNPFTAIGCAM 2203
         |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db  1141 LPPPPRSVPGVSCPESLQRSDPLEGPSXLPSSPPVLQLAMPMPLLGAGECNPFTAIGCAM 1200

Qy 2204 TETGGGPDDLPSYPPKKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPT 2263
         |||:|:  ||||||  ||||||||||||||:::||||||||||||||||||||||:||||
Db  1201 TETXGPXXLPSYPPKKEVSEWSDESWSTTTTASSVVTGPPYPKIRGKDSTQSATAKRPT 1260

Qy 2264 KKKLGKSEFSCCSMSITWTDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQ 2323
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 KKKLGKSEFSCCSMSITWTDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQ 1320

Qy 2324 KVTINRQLPFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVR 2383
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 KVTINRQLPFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTD-- 1378

Qy 2384 SGAARKAVLDLQKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLI 2434
```

```
Db   1261 KKKLGKSEFSCSMSYTWTDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQ 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2324 KVTINRQPLFPPSYHKQVRLAKEKASKVGVMDYDEVAAHTPSKSAKSHITGLRGTDVR 2383
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 KVTINRQPLFPPSYHKQVRLAKEKASKVGVMDYDEVAAHTPSKSAKSHITGLRGTD-- 1378
Qy   2384 SGAARKAVLDLQKCVEAGEIPSHYRQTVPKEEVFVKTPQKPTKKPPRLI 2434
                ||||||||||||||||||||||||||||||||||||||||||||
Db   1379 ------VLDLQKCVEAGEIPSHYRQTVPKEEVFVKTPQKPTKKPPRLI 1422
```

RESULT 7
US-08-467-344A-83
; Sequence 83, Application US/08467344A
; Patent No. 6586568
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
;   TAMI J. PILOT-MATIAS
;   GEORGE J. DAWSON
;   GEORGE G. SCHLAUDER
;   SURESH M. DESAI
;   THOMAS P. LEARY
;   ANTHONY SCOTT MUERHOFF
;   JAMES C. ERKER
;   SHERI L. BUIJK
;   ISA K. MUSHAHWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
;   REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES   D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,344A
; FILING DATE: 07-Jun-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/424,550
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: POREMBSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1422 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 83:
US-08-467-344A-83

Query Match 48.0%; Score 7380; DB 4; Length 1422;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1392; Conservative 2; Mismatches 27; Indels 10; Gaps 2;

```
Qy   1005 AHPTGSIHPITVDAANDDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYW 1064
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AHPTGSIHPITVDAANDDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYW 60
```

```
Db    181 TYGVNPNCYFNGKCTNTGASLTYSTYGMYLTGRCSRNYDVIICDECHATDRRTVLGIGKV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1245 LTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDEGTIPHGKKIKEENLKKGRHLI 1304
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 LTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDEGTIPHGKKIKEENLKKGRHLI 300
Qy   1305 FEATKKHCDELANELARKGITAVSYYRGCDISKIPEGDCVVATDALCTGYTGDFDSVYD 1364
          ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db    301 FEATKKHCDELANELARKGITAVSYYRGCDISKMPEGDCVVATDALCTGYTGDFDSVYD 360
Qy   1365 CSLMVEGTCHVDLDPFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVP 1424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CSLMVEGTCHVDLDPFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVP 420
Qy   1425 ECNIVEAFDAAKAWYGLSSTEAQTILDYTRTQPGLPAIGANLDEWADLFSMVNPEPSFVN 1484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ECNIVEAFDAAKAWYGLSSTEAQTILDYTRTQPGLPAIGANLDEWADLFSMVNPEPSFVN 480
Qy   1485 TAKRTADNYVLLTAAQLCHQYGVAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPE 1544
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db    481 TAKRTADNYVLLTAAQLCHQYGVAAPNDAPRWQGARLGKKPCGVLWRLDGCDACPGPE 540
Qy   1545 PSEVTRYQMCFTEVNTSGTAALAVGVAMAYLAIDFTGATCVRRCWSITSVPTGATVAP 1604
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 PSEVTRYQMCFTEVNTSGTAALAVGVAMAYLAIDFTGATCVRRCWSITSVPTGATVAP 600
Qy   1605 VVDEEEIVEECASFIPLEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAI 1664
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 VVDEEEIVEECASFIPLEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAI 660
Qy   1665 IEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMM 1724
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    661 IEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARXALAFMM 720
Qy   1725 AGAAGTALGTWTSVGFVDMLGGYAAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAA 1784
          ||| ||||||||||||||||||||| ||||||||||||||||| |||||||||||||||||
Db    721 AGAXGTALGTWTSVGFVDMLGGYAGASSTACLTFKCLMGEWXTMDQLAGLVYSAFNPAA 780
Qy   1785 GVVGVLSACAMPFALTTAGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTP 1844
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    781 GVVGVLSACAMPFALTTAGPDHWPNRLLTMLARSNTVCKEYFIATRDIRRKILGILEASTP 840
Qy   1845 WSVISACIRWLHTPTEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYS 1904
          || |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    841 WSXISACIRWLHTPTEDDCGLIAWGLXIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYS 900
Qy   1905 CQKGYKGPWIGSGMLQARCPCGAELIPSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGS 1964
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    901 CQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGS 960
Qy   1965 ARPDPTDWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDGQV 2024
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||  |||
Db    961 ARPDPTDWTSLVVNYGVRDYCKYEKLGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDRVQV 1020
Qy   2025 QCYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNN 2084
          | ||||||||||||||||||||||||||||||||||||| |||||||||| || ||||| ||
Db   1021 QXYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGGRMQLNLRDRLEANDCNSINN 1080
Qy   2085 TPSDEAAVSALVFKQELRRTNQLLEAISAGVDTTKLPAPS-IEEVVRKRQFRARTGSLT 2143
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   1081 TPSDEAAVSALVFKQELRRTNQLLEAISAGVDTTKLPAPSQIEEVVRKRQFRARTGSLT 1140
Qy   2144 LPPPPRSVPGVSCPESLQRSDPLEGPSNLPPSPPVLQLAMPMLLGAGECNPPTAIGCAM 2203
          |||||||||||||||||||||||||||| || ||||||||||||||||||||||||||||
Db   1141 LPPPPRSVPGVSCPESLQRSDPLEGPSXLPSSPPVLQLAMPMLLGAGECNPPTAIGCAM 1200
Qy   2204 TETGGGPDDLPSYPPKKEVSEWSDESWSTATTASSYYTGPPYPKIRGDSTQSAPAKRPT 2263
          |||  | || |||||||||||||||||||  |||||||||||||||||||||||| |||||
Db   1201 TETXGXPXXLPSYPPKKEVSEWSDESWSTTTTASSYYTGPPYPKIRGDSTQSATAKRPT 1260
Qy   2264 KKKLGKSEFSCSMSYTWTDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQ 2323
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy 1065 CVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAH 1124
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVGQIRVRPLVCAGYHPQYTAH 120

Qy 1125 ATLDTKPTVPNEYSVQILIAPTGSGKTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHA 1184
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ATLDTKPTVPNEYSVQILIAPTGSGKTKLPLSYMQXKXEVLVLNPSVATTASMPKYMHA 180

Qy 1185 TYGVNPNCYFNGKCTNTGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKV 1244
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 TYGVNPNCYFNGKCTNTGASLTYSTYGMYLTGRCSRNYDVIICDECHATDRTTVLGIGKV 240

Qy 1245 LTEAPSKNVRLVVLATATPGVIPTPHANTEIQLTDEGTIPFHGKKIKEENLKKGRHLI 1304
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 LTEAPSKNVRLVVLATATPGVIPTPHANTEIQLTDEGTIPFHGKKIKEENLKKGRHLI 300

Qy 1305 FEATKGHCDELANELARKGITAVSYRGCDISKIPEGDCVVATDALCTGYTGDFDSVYD 1364
        |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db  301 FEATKGHCDELANELARKGITAVSYRGCDISKMPEGDCVVATDALCTGYTGDFDSVYD 360

Qy 1365 CSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVP 1424
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 CSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVP 420

Qy 1425 ECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVN 1484
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 ECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVN 480

Qy 1485 TAKRTADNYVLLTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPE 1544
        |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  481 TAKRTADNYVLLTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGDACPGPE 540

Qy 1545 PSEVTRYQMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAP 1604
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 PSEVTRYQMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAP 600

Qy 1605 VVDEEIVEECASFIPLEAMVAAIDKLKSTITTSPFTLETALEKLNTFLGPHAATILAI 1664
        |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db  601 VVDEEIVEECASFIPLEAMVAAIDKLKSTITTSPFTLETALEKLNTFLGPHAATILAI 660

Qy 1665 IEYCCGLVTLPDNPASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMM 1724
        |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db  661 IEYCCGLVTLPDNPASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARXALAFMM 720

Qy 1725 AGAAGTALGTWTSVGFVFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAA 1784
        |||| |||||||||||||||||| |||||||||||||||||| |||||||||||||||||
Db  721 AGAXGTALGTWTSVGFVFDMLGGYAGASSTACLTFKCLMGEWXTMDQLAGLVYSAFNPAA 780

Qy 1785 GVVGVLSACAMFALTTAGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTP 1844
        |||||||||||||||||||||| |||||||||||||||| ||||||||||||||||||||
Db  781 GVVGVLSACAMFALTTAGPDIWPNRLLTMLARSNTVCXEYFIATRDIRRKILGILEASTP 840

Qy 1845 WSVISACIRWLHTPEDDCGLIAWGLEIWQVVCNFFVICFNVLKAGVQSMVNIPGCPFYS 1904
        || ||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  841 WSXISACIRWLHTPEDDCGLIAWGLXIWQVVCNFFVICFNVLKAGVQSMVNIPGCPFYS 900

Qy 1905 CQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGS 1964
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 CQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGS 960

Qy 1965 ARPDPTDWTSLVVNYGVRDYCKYBKMGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDGVQV 2024
        ||||||||||||||||||||||||| ||||||||||||||||||||||||||||| ||||
Db  961 ARPDPTDWTSLVVNYGVRDYCKYEKLGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDRVQV 1020

Qy 2025 QCYLGEPKTPWTTSACCYGPDGKGTVKLPPRVDGHTPGVRMQLNLRDALETNDCNSTNN 2084
        | |||||||||||||||||||||| ||||||||||||| |||||||||| |||||| ||
Db 1021 QXYLGEPKTPWTTSACCYGPDGKGKTVKLPPRVDGHTPGGRMQLNLRDRLEANDCNSINN 1080

Qy 2085 TPSDEAAVSALVFKQELRRTNQLLEAISAGVDTTTKLPAPS-IEEVVVRKRQFRARTGSLT 2143
        |||||||||||||||||||||||||||||||||||||||||| |||||||| |||||||||
Db 1081 TPSDEAAVSALVFKQELRRTNQLLEAISAGVDTTTKLPAPSQIEEVVVRKKQFRARTGSLT 1140

Qy 2144 LPPPPRSVPGVSCPESLQRSDPLEGPSNLPPSPPVLQAMPLLGAGECNPFTAIGCAM 2203
```

```
Db 1141 LPPPPRSVPGVSCPESLQRSDPLEGFSXLPSSPPVLQLAMPMPLLGAGECNPFTAIGCAM 1200

Qy 2204 TETGGGPDDLPSYPPKEVSEWSDESWSTATTASSVVTGPPYPKIRGKDSTQSAPAKRPT 2263
        |||  | ||||||| |||||||||||| ||||||| ||||||||||||||||| |||||
Db 1201 TETXGXPXXLPSYPPKKEVSEWSDESWSTTTTASSYVTGPPYPKIRGKDSTQSATAKRPT 1260

Qy 2264 KKKLGKSEFSCCSMSYTWTDVISFKTASKVLSATRAITSGFLKQRSLYVVTEPRDAELRKQ 2323
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 KKKLGKSEFSCCSMSYTWTDVISFKTASKVLSATRAITSGFLKQRSLYVVTEPRDAELRKQ 1320

Qy 2324 KVTINRQPLFPPSYHKQVRLAKEKASKVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVR 2383
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 KVTINRQPLFPPSYHKQVRLAKEKASKVGVMWDYDEVAAHTPSKSAKSHITGLRGTD-- 1378

Qy 2384 SGAARKAVLDLQKCVERAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLI 2434
              ||||||||||||||||||||||||||||||||||||||||||||
Db 1379 ------VLDLQKCVERAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLI 1422
```

RESULT 8
US-08-424-550B-83
; Sequence 83, Application US/08424550B
; Patent No. 6720166
GENERAL INFORMATION:
APPLICANT: JOHN N. SIMONS
APPLICANT: TAMI J. PILOT-MATIAS
APPLICANT: GEORGE J. DAWSON
APPLICANT: GEORGE G. SCHLAUDER
APPLICANT: SURESH M. DESAI
APPLICANT: THOMAS P. LEARY
APPLICANT: ANTHONY SCOTT MUERHOFF
APPLICANT: JAMES C. ERKER
APPLICANT: SHERI L. BUIJK
APPLICANT: ISA K. MUSHAHWAR
TITLE OF INVENTION: NON-A, NON-B. NON-C, NON-D, NON-E HEPATITIS
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
NUMBER OF SEQUENCES: 716
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES   D377/AP6D
STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,550B
FILING DATE:
CLASSIFICATION: 435435
ATTORNEY/AGENT INFORMATION:
NAME: POREMBSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5527.PC.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 1422 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-424-550B-83

Query Match 48.0%; Score 7380; DB 4; Length 1422;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1392; Conservative 2; Mismatches 27; Indels 10; Gaps 2;

```
Qy  2085  TPSDEAAVSALVFKQELRRTNQLLEAISAGVDTTKLPAPS-IEEVVRKRQFRARTGSLT  2143
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db  1081  TPSDEAAVSALVFKQELRRTNQLLEAISAGVDTTKLPAPSQIEEVVRKRQFRARTGSLT  1140

Qy  2144  LPPPPRSVPGVSCPESLQRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAM  2203
          |||||||||||||||||||||||||||| || ||||||||||||||||||||||||||||
Db  1141  LPPPPRSVPGVSCPESLQRSDPLEGPSXLPSSPPVLQLAMPMPLLGAGECNPFTAIGCAM  1200

Qy  2204  TETGGGPDDLPSYPPKKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPT  2263
          |||  | ||||||||||||||||||||| |||||||||||||||||||||||||| ||||
Db  1201  TETXGXPXXLPSYPPKKEVSEWSDESWSTTTTASSYVTGPPYPKIRGKDSTQSATAKRPT  1260

Qy  2264  KKKLGKSEFSCSMSYTWTDVISFKTASKVLSATRAITSGFLKQRSLVVYTEPRDAELRKQ  2323
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  KKKLGKSEFSCSMSYTWTDVISFKTASKVLSATRAITSGFLKQRSLVVYTEPRDAELRKQ  1320

Qy  2324  KVTINRQPLFPPSYHKQVRLAKEKASKVVGVMNDYDEVAAHTPSKSAKSHITGLRGTDVR  2383
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  KVTINRQPLFPPSHKQVRLAKEKASKVVGVMNDYDEVAAHTPSKSAKSHITGLRGTD--  1378

Qy  2384  SGAARKAVLDLQKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLI  2434
                ||||||||||||||||||||||||||||||||||||||||||||
Db  1379  ------VLDLQKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLI  1422
```

RESULT 9
US-07-925-695-8
; Sequence 8, Application US/07925695
; Patent No. 5428145
GENERAL INFORMATION:
  APPLICANT: OKAMOTO, Hiroaki
  APPLICANT: NAKAMURA, Tetsuo
  TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME,
  TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
  TITLE OF INVENTION: DETECTION SYSTEMS
  NUMBER OF SEQUENCES: 9
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Beveridge, DeGrandi, Weilacher & Young
    STREET: 1850 M Street, N.W., Suite 800
    CITY: Washington
    STATE: D.C.
    COUNTRY: US
    ZIP: 20036
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/07/925,695
    FILING DATE: 1920807
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: JP 287402/91
    FILING DATE: 09-AUG-1991
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: JP 360441/91
    FILING DATE: 05-DEC-1991
  ATTORNEY/AGENT INFORMATION:
    NAME: Weilacher, Robert G.
    REGISTRATION NUMBER: 20,531
    REFERENCE/DOCKET NUMBER: 06/87-48009
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (202) 659-2811
    TELEFAX: (202) 659-1462
    TELEX: WUI 64470
INFORMATION FOR SEQ ID NO: 8:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 3033 amino acids
    TYPE: AMINO ACID
    STRANDEDNESS: unknown
    TOPOLOGY: linear
US-07-925-695-8

```
Qy  1005  AHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYW  1064
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  AHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYW    60

Qy  1065  CVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAH  1124
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db    61  CVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVGQIRVRPLVCAGYHPQYTAH   120

Qy  1125  ATLDTKPTVPNEYSVQLIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHA  1184
          ||||||||||||||||||||||||||||||||||| | ||||||||||||||||||||
Db   121  ATLDTKPTVPNEYSVQLIAPTGSGKSTKLPLSYMQXKEVLVLNPSVATTASMPKYMHA   180

Qy  1185  TYGVNPNCYFNGKCTNTGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKV  1244
          |||||||||||||||||||||||||||||||| ||||||||||||||||| |||||||||
Db   181  TYGVNPNCYFNGKCTNTGASLTYSTYGMYLTGRCSRNYDVIICDECHATDRTTVLGIGKV   240

Qy  1245  LTEAPSKQNVRLVVLATATPPGVIPTPHANITEIQLTDEGTIPPHGKKIKEENLKKGRHLI  1304
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  LTEAPSKQNVRLVVLATATPPGVIPTPHANITEIQLTDEGTIPPHGKKIKEENLKKGRHLI   300

Qy  1305  FEATKKHCDELANELARKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYD  1364
          |||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db   301  FEATKKHCDELANELARKGITAVSYYRGCDISKMPEGDCVVVATDALCTGYTGDFDSVYD   360

Qy  1365  CSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGITYYYVDGSCTPSGMVP  1424
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  CSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGITYYYVDGSCTPSGMVP   420

Qy  1425  ECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVN  1484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  ECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVN   480

Qy  1485  TAKRTADNYVLLTAAQLCHQYGVYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPE  1544
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db   481  TAKRTADNYVLLTAAQLCHQYGVYAAPNDAPRWQGARLGKKPCGVLWRLDGCDACPGPE   540

Qy  1545  PSEVTRYQMCFTEVNTSGTAALAVGVGVAMAYLAIDFTGATCVRRCWSITSVPTGATVAP  1604
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  PSEVTRYQMCFTEVNTSGTAALAVGVGVAMAYLAIDFTGATCVRRCWSITSVPTGATVAP   600

Qy  1605  VVDEEEIVEECASFIPLEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAI  1664
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  VVDEEEIVEECASFIPLEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAI   660

Qy  1665  IEYCCGLVTLPDNPFASCVFAFIAGITTTPLPHKIKMFLSLFGGAIASKLTDARGALAFMK  1724
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db   661  IEYCCGLVTLPDNPFASCVFAFIAGITTTPLPHKIKMFLSLFGGAIASKLTDARXALAFMK   720

Qy  1725  AGAAGTALGTWTSVGFVFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAA  1784
          |||:||||||||||||||||||||| ||||||||||||||||:|||||||||||||||||
Db   721  AGAXGTALGTWTSVGFVFDMLGGYAGASSTACLTFKCLMGEWXTMDQLAGLVYSAFNPAA   780

Qy  1785  GVVGVLSACAMFALTTAGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKIIGILEASTP  1844
          |||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db   781  GVVGVLSACAMFALTTAGPDHWPNRLLTMLARSNTVCKEYFIATRDIRRKIIGILEASTP   840

Qy  1845  WSVISACIRWLHTPTEDDCGLIAWGLEIWQYVCNFVVCFNVLKAGVQSMVNIPGCPFYS  1904
          ||:|||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db   841  WSXISACIRWLHTPTEDDCGLIAWGLXIWQYVCNFVVCFNVLKAGVQSMVNIPGCPFYS   900

Qy  1905  CQKGYKGPWIGGSMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGS  1964
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  CQKGYKGPWIGGSMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGS   960

Qy  1965  ARPDPTDWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVADVGVQV  2024
          ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||:|||
Db   961  ARPDPTDWTSLVVNYGVRDYCKYEKLGDHIFVTAVSSPNVCFTQVPPTLRAAVADVRVQV  1020

Qy  2025  QCYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNN  2084
          |:|||||||||||||||||||||||||||||||||||||:||||||||:||:||||||:||
Db  1021  QXYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGGRMQLNLRDRLEANDCNSINN  1080
```

```
Query Match       21.3%;   Score 3274.5;  DB 1;  Length 3033;
Best Local Similarity  29.9%;   Pred. No. 1.9e-272;
Matches  980;  Conservative  437;  Mismatches 1181;  Indels  685;  Gaps 107;
```

```
Qy      8  TSPVPAPRTRKNKQTQASYPVSIK----------TSVERGQRAKRKVQRDARPR   51
           |:|| |  ||||      ::|      |:      ||||| :||:|  | :|
Db      3  TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVTLLPRRGPRLGVRATRKTSERSQPR   59

Qy     52  NYKIAGIHDGLQTLAQAALPAH-----------------GWGRQDPRHKSRNL         87
           || || |||  ::                          |:||||| ||:||||
Db     60  GRRQPIPKDRRSTGSWGKPGPYPWPLYGNEGCGWAGWLLSPRGSRPTWGPTDPRHRSRNL  119

Qy     88  GILLDYPLGWIGDVTTHTPLVGPLIVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLFV  144
           ||  | ||||| || :||| |:|||| ||| | | || ||||:| || |   ||| |||:
Db    120  GRVIDTITCGFADLMGYIPVVGAPV-GGVARALAHGVRVLEDGINYATGNLPGCSFSIFL  178

Qy    145  VCLLS-LACPCSGARVTDPDTNTTILTNCCRNQVIYCSPSTCLHEPGCVICADE------  198
           | |||  |  |      ||    ||: |  |   :|    |  |  ||||| |||
Db    179  LALLSCVTVPVSAVEVRN-ISSSYYATNDCSNNSITWQLITDAVLHLPGCVPCENDNGTLH  237

Qy    199  CWVPANPYISHPSNWTGTDSPLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV----  254
           ||:|     :   |  :|     :||| |||| |:|  ||| ::
Db    238  CWIQVTPNVAVKHRGALTRS-LRTHVDMIVMAATACSALYGDVCGAVMILSQAFMVSPQ   296

Qy    255  RHWLIHIDLNETGTCYLEVPTGIDEGFLGFIGWMAGKVEAVIFLTKLASQVPYAIATMFS  314
           || |     |  :|     ||    |:   |  |    : :|| ||||||  ||:::|
Db    297  RH------NFTQECNCSIYQGHITGHRMAWDMMLSWSPTLTMILAYAARVPELVLEIIF   349

Qy    315  SVHYLAVGALIYYASRGKWYQLLLAIMLY--IEAT-----SGNPIRVPTGCSIAEFCSPLMI  369
           ||| |||  | |   |:  ||  | |      |:      :| |:||| || |  |:|:
Db    350  GGHWGVVFGLAYFSMQGAWAKVIAILLLVAGVDATTYSSGQ---EAGRTVAGFAGLFTT   405

Qy    370  PCP---------------------CHSYLSENVSEVICYSPKWT-----------RPI---  395
           |||                     ||: :  || ||  || |             ||:
Db    406  GAKQNLYLINTNGSWHINRTALNCNDSLQTGFLASLFYTHKFNSSSGCPERLSSCRGLDDF  465

Qy    396  ---------TLEYNNSIS----------W-YP--------YTIPGARGCM--------   417
              ::        | :             | ||        || | | |
Db    466  RIGWGTLEYETNVTNDGDMRPYCWHYPRPCGIVPARTVCGPVYCFTPSPVVVGTTDKQG  525

Qy    418  ---------VKFKNNT---------WGC--------------CRIR----NVPSYC----  437
                    |||:|:          |||              ||:|     ||| |
Db    526  VPTYTWGENETDVFLLNSTRPPRGAWFGCTWMNGTGFTKTCGAPPCRIRKDYNSTIDLLC  585

Qy    438  TMGTDAVWNDTRNTYEACGVTPWLT--------TAWHNGSALKLAILQ----YPGSKEMF  485
           | |   |  ||  ||    | ||:|         || |  |  | |:     ||||:|:
Db    586  P---TDCFRKHPDATYLKCGAGPWLTPRCLVDYPYRLWHYPCTVNFTIFKARMYVGGVE-  641

Qy    486  KPHNWMSGHLYF------------EGSDTPIVYFYDPVNSTLLPPERWARLPGTPVVRGS  534
           ||| | |  | |            | | |:  | || :     | :       | |
Db    642  --HRFSAACNFTRGDRCRLEDRDRGQQSPLLH------ST---TEWAVLPCS--------  682

Qy    535  WLQVPQGFYSDVKDLATGLLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSK  594
           || |  |||      | |:|      :|||::|:|  : ||:||:       |:  | |
Db    683  ------FSDLPALSTGLLHLHQNIVDVQYLYGLSPALTRYIVKNEWVILLFLLIADAR   734

Qy    595  YLILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYLRCR  654
           :|:|  ||     |||  | |:|| |   |    | |:| | ||  | ||: |  | |||
Db    735  --ICACLWMLI-------------------ILGQAEAALEKLII--------------   757

Qy    655  LRYAAALLGFVEPMAAGLPLTFFVAAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALLVGP  714
           |||   || | |   |:|   || | |  || | | :|    | | |||||   :| | | |
Db    758  -----LHSASAASANGPLWFIFFTAA---WYLKGRVUPVATYS------------VLGL   797

Qy    715  WPLVALLTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLW-  773
           | |:| | ||:|           |: |||   |    | |     ||| | | |||
Db    798  WSF-LLLVLALPQQAYALDRAAEQGEGLGLAILVIISIFTLTPAYKILLSRSVWWLSYMLV  855

Qy    774  -----QRW---------ENWFW-NVTLRPERFFLV---LVCFPGATYDALVTFCVCHV   813
                ||         |||:|  ||:  || |:|    |:||| || |  | | | ||:
Db    856  LAEAQIQQWWPPLEVRGGRDGIIWVAVILHPRLVFEVTKWLLAILGPAY------AW  904

Qy    814  ALLCLTSSAASFFGTDSRVRAHRML--VRLGKCHAWYSHYVLKFFLLVFGENGVFFYGHL  871
           ||| | |        | : : |  |   ||| : | :    |  ::| | | |     |
Db    905  -LLKASLRIPYF-----VRAHALLRVCTLVK-HLAGARYIQMLLITIGRWTGTYIYDHL  957

Qy    872  HGDVLPNDFASKLPLQ----------EP--FPPEGKARVYRNEGRRLACGDTVDG      915
               | |  | |:|||          ||  |||||| || | ||| ||| ||
Db    958  S-------------PLSTWAAQGLRDLAIAVEPVVFSMEKKVIVWGAE--TVACGDILHG  1003

Qy    916  LPVVARLGDLVFAGLAMPPD-----GWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWT  970
           ||| ||||| : |||        | |   |   ||   :|||||  ||  | | |:||
Db   1004  LPVSARLGREVLLG----PADGYTSKGWKLLAPITAYTQQTRGLLGAIVVSLTGRDKNEQA  1060

Qy    971  GTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDIYQPPCG  1030
           | |:| |: ::   | |:|:| |||:|| ||: ||| || : |  || |    : |  |
Db   1061  GQVQVLSSVTQTFLGTSISGVLWTVYHGAGNKTLAGPKGPVTQMYTSAEGDLVGWPSPPG  1120

Qy   1031  AGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPLLCSSGHV  1090
           ||| | || | ||| || | || || :| | ||| |||||  | | |||||| ||  |
Db   1121  TKSLDPCTCGAVDLYLVTRNADVIPVRRKDDRRGALLSPRPLSTLKGSSGGPVLCSRGHA  1180

Qy   1091  ICGMFTA----ARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILIAPTG  1147
           | |  |:    ||:  |  |    ||| |:      ||  ::|  | | :  |:   | |||
Db   1181  VGLFRAAVCARGVAKSIDFIPVESLDVATRTPSFSDNS---TPPAVPQSYQVGYLHAPTG   1237

Qy   1148  SGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCNTGASLTY  1207
           |||||  ||||| |||||||||||| | |  ||| |||| || |  :|:  |   | ||
Db   1238  SGKSTKVPAAYASQGYKVLVLNPSVAATLGFGAYMSKAHGINPNIRTGVRTVTTGDSITY  1297

Qy   1208  STYGMYLT-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPG  1265
           |||| ||  || | | ||| ||||| ||  | :||||| | |   ||||| |||||||||
Db   1298  STYGKFIADGGCAAGAYDIIICDECHSVDATTILGIGTVLGQAETAGVRLVLLATATPPG  1357

Qy   1266  VIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELARKGIT  1325
           | ||| |:|:| ::| | | | ||: : | |: || ||||| :| | |||| ||  |  |
Db   1358  TVTTPHSNIEEVALGHEGEIPFYGKAIPLAFIKGGRHLIFCHSKKKCDELAAALRGMGVN  1417

Qy   1326  AVSYYRGCDISKIP-EGDCVVVATADLCTGYTGDFDSVYDGSLMVEGTCHVDLDPTFTMG  1384
           |  ||||||| ||  |||| ||:|| | |||||||| ||| | |: | || |||| |   |
Db   1418  AVAYYRGLDVSVIPTQGDVVVVATDALMTGYTGDFDSVIDCNVAVSQVDFSLDPTFTIT  1477

Qy   1385  VRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAPFDAAKAWYGLSST  1444
           | |||| ::|  |||| ||||| ||||:||||| |||| | | |||||  || |  |:  |
Db   1478  TQTVPQDAVSRSQRRGRTGRGRLGVYRYVSSGERPSGMFDSVVLCECYDAGAAWYELTPA  1537

Qy   1445  EAQTILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLQL  1503
           | ||:||  |   |||  ::  ||   || |:| |||:|  :|  ||: |  |   |  |
Db   1538  ETTVRLRAYFNTPGLPVCQDHLEFWEAVFTGLTHIDAHFLSQTKQGGENFAYLTAYQATV  1597

Qy   1504  CHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADAC-----------PQPEP-----  1545
           |  |||| |    |  | ::     | :| | | |  |           | ||
Db   1598  CAR----AKAPPPSWD----------VMWK-------------CLTRLKPTLTGPTLLYRLGAVTN  1637

Qy   1546  --------------SEVTRY-----QMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITS-  1595
                         |:|  |     |:|  ||  | ||  |:| | || | : ||| |    |
Db   1638  EVTLTHPVTKYIATCMQADLEIMTS-SWVLAGGVLAAAYCLAT------GCISIGR  1689

Qy   1596  --VPTGATVAPVVDEE-------EIVECASFIPL-EAMVAAIDKLKSTI------TTTSP  1640
             || |  |   ||       | :|||| | |  | |:  ||: |||
Db   1690  LHLNDRVVVAP--DKEILYEAFDEMEECASKAALIEBGQRMAEMLKSKIQGLLQQATRQA  1747

Qy   1641  FTLETALE----KLNTFLGPHAATILAIIEYCCGLVTLPDNPFASCVFAFIAGITTTPLPH  1696
           | | |||     |:| |||   |||  |:||| ||||| |||  |  | | |:| ||||
Db   1748  QDIQPAIQSSWPKLEQFWAKHMWNFISGIQYLAGLSTLPGNPAVASMMAFSAALTSPLPT  1807

Qy   1697  KIKMFLSLFGGAIASKLTDARGALAFWMAGAAGTALGTWTSVG---FVFDMLGGYAAASS  1753
           | :||||| |||  | |  |  | |  |:  | |  |  | | |   ::|| || || |:|
Db   1808  STTILLNIMGGWLASQIAPPAGATGFVVSGLVGAAVG---SIGLGKILVDVLAGYGAGIS  1864

Qy   1754  TACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTTAGPD--HWPNRLL  1811
           ||| ||| |||| || || |||:| |:|  | ||| ||| |:|| || | |   | | |
Db   1865  GALVAFKIMSGEKPTVEDVVNLLPAILSPGALVVGVICAAILRRHVGQGEGAVQWMNRLI  1924

Qy   1812  TMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPEDDCGLI----AW  1868
           | |||:| | ||   ||  |||| || : | |||  |: |  :||:|| ||
Db   1812  ...
```

```
Db   2937 PPLRAWKSRARAVRASLIAQGARAAICGRYLFNWAVKTKLKLTPLPEASRLDLSGWFT-- 2994
             :   |   |||
Qy   2821 YCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAI 2863
             |  ||:||:|||:|:|:|
Db   2995 ---VGAGGGDIYHSVSHARPRLL------LLCLLLSVGVGI 3027
```

RESULT 10
US-07-925-695-5
; Sequence 5, Application US/07925695
; Patent No. 5428145
; GENERAL INFORMATION:
; APPLICANT: OKAMOTO, Hiroaki
; APPLICANT: NAKAMURA, Tetsuo
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME,
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
; TITLE OF INVENTION: DETECTION SYSTEMS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weilacher & Young
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: US
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07,925,695
; FILING DATE: 19920807
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 287402/91
; FILING DATE: 09-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 360441/91
; FILING DATE: 05-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Weilacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/87-48009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; TELEX: WUI 64470
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3033 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: linear
US-07-925-695-5

Query Match   21.3%; Score 3269; DB 1; Length 3033;
Best Local Similarity   30.1%; Pred. No. 5.8e-272;
Matches 981; Conservative 432; Mismatches 1204; Indels 640; Gaps 112;

Qy   8 TSPVPAPRTRKNKQTQASYPVSIK------------TSVERGQRAKRKVQRDARPR 51
            :|:| ||::| | | :   :::|
Db   3 TNPKPQRKTKRNTNRR---PQDVKFPGGGOIVGGVYLLPRRGPRLGVRATRKTSERSQPR 59

Qy   52 NYKIAGIHDGLQTLAQALPAH----------GWGRQDPRHKSRNL 87
            ::|:| |||: | | |:| ::: |:|
Db   60 GRRQPIPKDRRSTGCKSWGKPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPNDPRHRSRNV 119

Qy   88 GILLDYPLGWIGDVTHTPLVG-PLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLF 143
           ||| | |||| : | |: | | | | ||| : | |:|
Db   120 GKVIDTLTCGFADLMGYIPVVGAPL--GVARALAHGVRVLEDGVNFATGNLPGCSFSIF 177

Qy   144 VVCLLS-LACPCSGARVIDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC-----AD 197
           :||| : || | | : | || | || : | | | |:|
```

Db  1137  VTRNADVIPARRRGDKRGALLSPRPLSTLKGSSGGPVLCPRGHAVGVFRAAVCSRGVAKS  1196
          ::||||||  | || | |  | |    ||||   ||||||:|||   |:   | || ||
Qy  1104  VSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQEKY  1163

Db  1197  IDFIPVETLDIVTRSPTFSDNS---TPPAVPQYQVYLHAPTGSGKSTKVPVAYAAQGY  1253
          : |:||::|:|  |  ||   ||     |||:|| |||||||||||| |:|||  |||
Qy  1164  EVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCNTGASLTYSYGMYLT-GACSRN-  1221

Db  1254  KVLVLNPSVAATLGFGAYLSKAHGINPNIRTGVRTVTGAPTYSTYGKFLADGGCAGGA  1313
          ||||||||| || |  |  |:|  | |   || |: || | : || |    |  || 
Qy  1222  YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTD  1281

Db  1314  YDIIICDECHAVDSTTILGIGTVLDQAETAGVRLTVLATATPPGSVTTPHPNIEEVALGQ  1373
          |||||||||| |:||:|||||:||| ||||||||||||||||| | ||| || |   | 
Qy  1282  EGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELARKGITAVSYYRGCDISKIP-E  1340

Db  1374  EGEIPFYGRAIPLSYIKGGRHLIFCHSKMKCDELAAALRGMLNAVAYYRGLDVSVIPTQ  1433
          ||||||:| || ||||||||||||| |: |||||||| | ::|| ||||| |||||| 
Qy  1341  GDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRG  1400

Db  1434  GDVVVVATDALMTGFTGDFDSVIDCNVATQVVDFSLDPTFITTQTVPQDAVSRSQRRG  1493
          || |||||||| ||:|||||||||| | || |||||| |:  ||  |  | |:|||||
Qy  1401  RTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLP  1460

Db  1494  RTGRGRLGIYRYVVSTGERASGMFDSVVLCECYDAGAAWYELTPAETTVRLRAYFNTPGLP  1553
          |||||| ||| | |||||||||| |||| ||| ||| |||| |||| ||| ||:|  |||
Qy  1461  AIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLCHQYGYAAPNDADPRWQ  1519

Db  1554  VCQDHLBFWEAVFTGLTHIDAHFLSQTKQSGENFAYLTAYQATVCAR----AKAPPSWD  1609
          ||   | |||   | |  |: |:||| :| |||||||||| | || |     || | ||
Qy  1520  GARLGKKPCGVLWRLDGADAC------------------------PGPEP---------Q  1552

Db  1610  ------------VMWK------CLTRLKPWLVGPTPLLYRLGSVTNEVTLTHPVTKYIATCM  1653
                      ||||      |||| ||| ||| ||||||  || || ||| ||| || |
Qy  1553  MCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI-  1611

Db  1654  QADLEWMTS-TWVLAGGVLAAVAAYCLAT-GCVCI----IGRLHVNQRAVVAP--DKEVLY  1706
          || ||||| |  ||| || || | |  || | ||    | || |||   |||  || |||
Qy  1612  -----VEECASFIPL-EAMVAAIDKLKSTI-------------TTTSPFTLETALE----  1655

Db  1707  EAFDEMEECASRAALIBEGQRIAEMLKSKIQGLLQQASKQAQDIQPAVQASWPKVEQFWA  1766
          ||||  ||||| || | ||||  || |::| |||| |||| | |  | ||| |||| |
Qy  1656  PHAATILAIIEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTD  1715

Db  1767  KHMWNFISGIQYLAGLSTLPGNPAVASMMAFSAALTSPLSTSTTILLNILGGWLASQTAP  1826
          | ||| |||||||||| ||| ||| ||||| ||| ||| |||| |||||| ||||| ||
Qy  1716  ARGALAFMMAGAAGTALGTWTSVGF----VFDMLGGYAAASTACLTFKCLMGEWPTMDQL  1772

Db  1827  PAGATGFVVSGLVGAAVG---SIGLGKVLVDILAGYGAGISGALVAFKIMSGEKPSMEDV  1883
          ||| |||| || |  |||   | || |||||||||| || ||| | | | || || |||
Qy  1773  AGLVYSAFNPAAGVVGVLSACAMFALTTAGPD----HWPNRLLTMLARSNTVCNEYFIAT  1828

Db  1884  VNLLPGILSPGALVVGVI--CAAILRRHVGPGEGAVQWMNRLLIAPASRGNHVAPTHYVTE  1941
          ||||| | || | ||||   |||| |||||| || | |||| | | ||||| |||| || 
Qy  1829  RDIRRKILGILEASTPWSVISACIRWLLHTPTEDDCGLI---AWGLEIWQYVCNFFVICPN  1885

Db  1942  SDASQRVTQLLGSLT----ITSLLRRLHNWITEDCPIPCSGSWLRDVWDWVCTILTDFKN  1997
          | ||||| |||| |     |||| || ||  |||||||| | |||||||||| ||||| |
Qy  1886  VLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPR  1945

Db  1998  WLTS--KLFPKMPGLPFISCQKGYKGVWAGTGIMTTRCPCGANISGNVRLGSMRI-TGPK  2054
          | ||  ||||||| |||||||||||| || | || |||||||| ||||||||| | ||
Qy  1946  TCSNYWRGAVPVPNARLCGSARPDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV  2004

Db  2055  TCMNIWQGTFPINCYTBGQCVPKPAPNFKAIAIWRVAASEYAEVTQHGSYHYIIGLTTDNL  2114
          ||||| || |||||| ||||||||| |||||||||| ||| |:|:|| || | |||||:|
Qy  2005  CFTQVP---PTLRAAVADVGVQVCYLGEPKTPWTTS-ACCYG----------------PD  2045

Db  2115  ---KVPCQLPSPEFFSWVDGVQIHRFAPIPKPFFFRDEVSFCVGLNSFVVGSQLPCDPEPD  2171
          :  || | | | |  |||| ||    |||| |||  ||| | ||| | ||| |  ||
Qy  2046  GKGKTVKL--PFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQBLLRR  2103

Db  2172  TDVLTSMLTDPSHITAETAARRL--------------ARGSPPSEASSSASQLSAPSLRA  2217

```
;   EARLIER FILING DATE: 1999-04-03
;   NUMBER OF SEQ ID NOS: 51
;   SOFTWARE: PatentIn Ver. 2.1
;
;   SEQ ID NO 3
;   LENGTH: 3010
;   TYPE: PRT
;   ORGANISM: Hepatitis C virus
US-09-539-601-3

Query Match       21.2%;   Score 3262;   DB 4;   Length 3010;
Best Local Similarity   29.8%;   Pred. No. 2.3e-271;
Matches  970;  Conservative  452;  Mismatches 1183;  Indels 648;  Gaps  116;
```

```
Qy      8  TSPVPAPRTRKNKQTQASYPVSIK----------------TSVERGQRAKRKVQRDARPR    51
           :|:|                                    : :||
Db      3  TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRPLGVRATRKTSERSQPR    59

Qy     52  NYKIAGIHDGLQTLAQAALP--------------------------------AHGWGRQD    79
           :  |: :|  | :
Db     60  GRR----------QPIPKARQPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPTD   111

Qy     80  PRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNATG---   136
           || :|||||||  ::|   |||| :
Db    112  PRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLP   170

Qy    137  WFGVHLFVVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC   195
           : |:||||||  |||  | ||||                    || |||| |||| ||
Db    171  GCSFSIFLLALLSCLTIPASAYEVRNV-SGVYHVTNDCSNASIYEAADMIMHTPGCVPC   229

Qy    196  ADE---------CWVPANPYISHPSNWTGTDSFPLADHIDFVMGALVTCDALDIGELCGACVLVG   250
           |||         :  |  |  || :::|
Db    230  VRENNSSRCWVALTPTIA-ARNASVPTTTIRRHVDLLVGAAALCSAMYVGDLCGSVFLVA   288

Qy    251  DWLV----RHWLIHIDLNETGTCYLEVPTGIDPGFL--GFIGW---MAGKVEAVIFLTKL   301
           ||||     :| :       |   : |         ::
Db    289  QLFTFSPRRHETVQ--------CSIYPGHVTGHRMAWDMMMWSPTAALVVSQL   337

Qy    302  ASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSIA   361
           | :|  |  ||::|| ||||||||||::          ||| ||  |||| | | | |
Db    338  L-RIPQAVVDMVAGAHWGVLAGLAYYSMVGNWAKVLIVMLLFAGVDGGTYV--TGGTMA   393

Qy    362  E----------------------------------------FCSPLMIPCPCHSYLSE   379
           :
Db    394  KNTLGITSLFSPGSSQKIQLVNTNGSWHINRTALNCNDSLNTGFLAALFY---VHKFNSS   450

Qy    380  -----NVSEVICYSPKWTRPITLEYNNSISWYPY------------------TIPGARGC   416
                 ||: |  | :|
Db    451  GCPERMASCSPIDAFAQGW-GPITYNESHSSDQRPYCWHYAPRPCGIVPAAQVCGPVYCF   509

Qy    417  -----------------------------------MVKFKNNT-----WGCCRIRNVPSY---C   437
                                              ||||:|       |||  ||||| |    |
Db    510  TSPVVVGTTDRFGVPTYSWGENETDVLLLNNTRPPQGNWFGCTWMNSTGFTKTCGGPPC   569

Qy    438  TMG-------------TDAVNDTRNTYEACCVTPWLT----------TAWHNGSALKLAILQ-   477
               :           :|  |  |        |||          :|
Db    570  NIGIGNKTLTCPTDCFRKHPEATYTKCGSGPWLTPRCLVHYPYRLWHYPCTVNFTIFKV   629

Qy    478  --YPGSKE--MFKPHNWMSG--HLYFEGSDTPIVVFYDPVNSTLLPPERWARLPGTPPVVR   532
             ::|            |           :  : ||
Db    630  RMVYGGVEHRLEAACNWTRGERCNLEDRDR------SELSPLLLSTTEWQVLPCS-----   678

Qy    533  GSWLQVPGQFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCG   592
                                               || :  ||  :||| |||  | :
Db    679  ----------------------------------FTTLPALSTGLIHLHQNVVDVQYLYGIGSAVVSFAIKWEYVLLLFLLLAD   728

Qy    593  SKYLILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPFALIFFICCYLR   652
           : :|             :  ||
Db    729  ARVCACLWMLL-----------------------IAQBAAALENLVV----------   753

Qy    653  CRLRYAALLGFVPMAAGLPLTFFVAAAAAQPDYDWWVR-LLIVAGLVLWAGRNRGHRIALL   711
               : ||                          : :         ||:   |
Db    754  --LNAASVAG----AHGI-LSFLVFFCAA-------WYIKGRLVPGAAY---------AL   790
```

```
Qy   2104  T--------NQLLEA-ISAGVDTTTKLPAPSIEEVVRKRQFRARTGSLTLPPPPRSVPG   2153
           :         ::|| | :|         ||
Db   2218  TCTTHGKAYDVDMVDANLFMGGDVTRIESES--KVVV-                       2252

Qy   2154  VSCPESLQRSDPL-EGPSNLPPS-----------PPVLQL-AMP---MPLLGA---GEC   2193
           :::                               :|  |  |:     | ||
Db   2253  ----LDSLDDMVEERSDLEPSIPSEYMLPKKRFPPALPAWARPDYNPPLVESWKRPDY   2306

Qy   2194  NPFTAIGCAMTETGGGPDDLPSYPPKKE----VSEWS--------------DES   2229
           :|:       : ||                : :|            ||
Db   2307  QPATVAGCALPP----PKKTPTPPPRRRRTVGLSESSIADALQQLAIKSFGCQPPPSGDSG   2362

Qy   2230  WST-ATTASSYVTGPP----------YPKIRGKDSTQSAPAKRPTKKKL---------   2267
               ::                     :||
Db   2363  LSTGADAADSGSRTPPDELALSETGSISSMPPLEGEPGD---PDLEPEQVELQPPPQGGV   2419

Qy   2268  ----GKSEFS--------CSMSYTWTD-VISFKTASKVLSATRAITSGFLKQRSLV   2310
               |:|||          | ||  | |  :|
Db   2420  VTPGSSGSGGWSTCSEEDDSVVCCSMSYSWTGALITPCSPEEEKLPINPLSNSLLRYHNKV   2479

Qy   2311  YVTEPRDAELRKQKVTINRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSA   2370
           |   |  |   ||  :|     |  :|    |  :  || |  :  |  |
Db   2480  YCTTSKSASLRAKKVTFDRMQALDAHYDSVLKDIKLAASKVTARLLTLEEACQLTPHSA   2539

Qy   2371  KSHITGLRGTDVR--SGAARKAVLDLQK-CVEAGEIPSHYRQTVIVPKEEVFVKTPQKPT   2427
           |:  |       |   |||    :|  ::|                 :|
Db   2540  RSKY-GFGAKEVRSLSGRAVNHIKSVWKDLLEDTQTPI---PTTIMAKNEVFCVDPTKGG   2595

Qy   2428  KKPPRRLISYPHLEMRCVEKMYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMWS--PD   2484
           |||| :|    ||||  :|:|         |:                | ||  ||  :
Db   2596  KKAARLIVYPDLGVRVCEKMALYDITQKLPQAVMGASYGFQYSPAQRVEFLLKAWAEKKD   2655

Qy   2485  AVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGRE   2544
           | | ||: ||| ||||      |  | |  | || :     || ||   | :||
Db   2656  PMGFSYDTRCFDSTVTERDIRTEESIYRACSLPEEAHTAIHSLTERLYVGGPMFNSKGQT   2715

Qy   2545  IGYRRCRSSGVVTTSSSNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADK   2604
           :|||||| | |         | ||  | |:| ||  |  : |||||  :  |  |
Db   2716  CGYRRCRASGVLTTSMGNTITCYVKALAAEQAGMRNCGTGILVCGDDLVVISESQGTEEDE   2775

Qy   2605  QAMRVFASWMKVMGAPQDCVPQPKYSLEELTCSSSNVTSGITKSGKPYYFLTRDPRIPLG   2664
           :||| |   |   | | ||   ||| |  || |  :::|  :||::|| |:|||  || |
Db   2776  RNLRAFTEAMTRYSAPPGDPPRPEYDLELITSCSSNVSALGPQGRRRYYLTRDPTTPIA   2835

Qy   2665  RCSAEGLGYNPSAAWIGYLIHHYPCLMVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYT   2724
           |  | |   |:|                |  :|         |        |||  :  |
Db   2836  RAAWETVRHSPVNSWLGNIIQVAPTIWARMVLMTHFFSILMAQDTLDQNLNFEMYGAVYS   2895

Qy   2725  VPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKR   2784
           :| |  |                                ||||     :|||: |
Db   2896  VSPLDLPAIIERLHGLDAFSLHTYTPHELTRVASALRKLGAPPLRAWKSRARAVRASLIS   2955

Qy   2785  RGGAHAKLARFLL-WHATSR----PLPDLDKTSVARYTTFNVCDVYSPEGDVFITPQRRL   2839
           |||                      :    ||: |  ||:|||::||   |||
Db   2956  RGGRAAVCGRYLFNWAVKTKLKLTPLPEARLLDLSSWFT-----VGAGGGDIYHSVSRAR   3010

Qy   2840  QKFLVKYLAVIVPALGL   2856
           ||| | :| |  |  ||
Db   3011  PRLLLLGLLLLFVGVGL   3027
```

```
RESULT 11
US-09-539-601-3
;   Sequence 3, Application US/09539601C
;   Patent No. 6630343
;   GENERAL INFORMATION:
;   APPLICANT: Bartenschlager, Ralf FW
;   TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
;   FILE REFERENCE: all sequences
;   CURRENT APPLICATION NUMBER: US/09/539,601C
;   CURRENT FILING DATE: 2001-08-30
;   EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
```

```
Qy   712 VGPWPLVALLTLLHLVTPASAFDTEI---IGGLTIPPVVALVVMSRPGFFAHLLPRCALV 768
Db   791 YGVWPL--LLLLLALPPRAYAMDREMAASCCGAVFVGLILLTLSPHYKLFLARL------- 842

Qy   769 NSYLWQRWENWFWNVTLRPERRFFLVLVCFP------GATYDALVTFCVCHVALL-CLTSSA 822
Db   843 ---IW----WLQYFITRABAHLQVWI--PPLNVRGGRDAVILLTCAIHPELIFTITKIL 892

Qy   823 ASFFG-------TDSRVRAHR------MLVRLGKCHAWYSHYVLKFFLLVFGENGVF 866
Db   893 LAILGPLMVLQAGITKVPYFVRAHGLIRACMLVR----KVAGGHVYQMALMKLAALTGTY 948

Qy   867 FYKHL-------HGDVLPNDFASKLPLQEPFPFEGKARVYRNEGRRLACGDTVDGLPVV 919
Db   949 VYDHLTPLRDWAHAGL--RDLAVAV---EPVVFSDMETKVITWGADTAACGDIILGLPVS 1003

Qy   920 ARLGDLVFAGLA--MPPDGWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLG 977
Db  1004 ARRGREIHLGPADSLEGQGWRLLAPITAYSQQTRGLLGCIITSLTGRDRNQVEGEVQVVS 1063

Qy   978 SLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN-DQDI--YQPPCGAGSL 1034
Db  1064 TATQSFLATCVNGVCWNTVYHGAGSKTLAGPKG---PITQMYTNVDQDLVGWQAPFGARSL 1120

Qy  1035 TRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMF 1094
Db  1121 TPCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSSGPLLCPSGHAVGIF 1180

Qy  1095 TAA----RNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILIAPTGSGKS 1151
Db  1181 RAAVCTRGVAKAVDFVPVESMETTMRSPVFTDNS---SPPAVPQTFQVAHLHAPTCSGKS 1237

Qy  1152 TKLPLSYMQEKYEVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYG 1211
Db  1238 TKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGAPITYSTYG 1297

Qy  1212 MYLT-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPT 1269
Db  1298 KFLADGGCSGGAYDIIICDECHSTDSTTILGIGTVLDQAETAGARLVVLATATPPGSVTV 1357

Qy  1270 PHANTEIQLTDEGTIPFHGKKIKEENLKKGRHLIPEATKKHCDELANELARKGITAVSY 1329
Db  1358 PHPNIEEVALSSTGEIPFYGKAIPIETIKGGRHLIFCHSKKKCDELAAKLSGLGLNAVAY 1417

Qy  1330 YRGCDISKIP-EGDCVVVATDALCTGYTGDFDSVVDCSLMVEGTCHVDLDPTFTMGVRVC 1388
Db  1418 YRGLDVSVIPTSGDVIVVATDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTV 1477

Qy  1389 GVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAPDAAKAWYGLSSTEAQT 1448
Db  1478 PQDAVSRSQRRGRTGRGRMGIYRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAETSV 1537

Qy  1449 ILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNVLLTAAQLCHQY 1507
Db  1538 RLRAYLNTPGLPVCQDHLEFWESVFTGLTHIDAHFLSQTKQAGDNFPYLVAYQATVCARA 1597

Qy  1508 GYAAPNDAPRWQG-ARLGKK---PCGVLWRLDGADACPGPEPSEVTRYQM-CFT---EVN 1559
Db  1598 QAPPPSWDQMWKCLIRLKPTLHGPTPLLYRL-GAVQNEVTTTHPITKYIMACMSADLEVV 1656

Qy  1560 TSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI------VE 1613
Db  1657 TS-TWVLVGGVLAALAAYCLTTGSVVIVGR------IILSGKPAIIPDREVLYREFDEME 1709

Qy  1614 ECASFIP-LEAMVAAIDKLK------STIT----TTSPFTLETALBKLNTFLGPHAATI 1661
Db  1710 ECASHLPYIEQCMQLAEQFKQKAIGLLQTATKQAEAAAP-VVESKWRTLEAFWAKHMWNF 1768

Qy  1662 LAIIEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIWKMPLSLFGGAIASKLTDARGALA 1721
Db  1769 ISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTTQHTLLFNILGGWVAQLAPPSAASA 1828

Qy  1722 FMMAGAAGTALGTWTSVGF---VFDMLGGYAAASSTACLITFKCLMGEWPTMDQLAGLVYS 1778


Db  1829 FVGAGIAGAAVG---SIGLGKVLVDILAGYGAGVAGALVAFKVMSGEMPSTEBDLVNLLPA 1885

Qy  1779 AFNPAAGVGVGVLSACAMFALTTAGPD----HWPNRLLTMLARSNTVCNEYFIATRDIRRK 1834
Db  1886 ILSPGALVVGVV--CAAILRRHVGPGEGAVQWMNRLLAFASRGNHVSPTHYVPESDAAAR 1943

Qy  1835 IIGILEASTPWSVISACIRWLHTPTEDDCGLI---AWGLEIWQYVCNFFVICFNVLKAGV 1891
Db  1944 VTQILSSLT----ITQLLKRLHQWINEDCSTPCSGSWLRDVWDWICT----VLTDFKTWL 1995

Qy  1892 QS--MVNIPGCPYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSN 1949
Db  1996 QSKLLPRLPGVPFFSCCQRGYKGWWRGDGIMQTTCPCGAQITGHVKNGSMRIV-GPRTCSN 2054

Qy  1950 YWRGAVPVNARLCGSARPDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFT 2007
Db  2055 TWHGTFPINAYTTGPCTPSPAPNYSRALWRVAAEBYVEVTRVGDFHYVTGMTTDNVKCPC 2114

Qy  2008 QVP------------------------VDGVQVQCYLGEPKTPWTTS 2038
Db  2115 QVPAPEFFTEVDGVRLHRYAPACKPLLREEVTFLVGLNQYLVGSQLPC---EPE------- 2165

Qy  2039 ACCYGPDGKGKTVKL--PFRVDGHTPGVRMQLNLRDALETND------------------- 2084
Db  2166 -----PDVAVLTSMLTDPSHITAETAKRLARGSPPSLASSSASQLSAPSLKATCTTRHD 2220

Qy  2085 TPSDEAAVSALVFKQEL------------RRTNQLL----EAISAGVDTTKLPAPSIEEVVRK 2132
Db  2221 SPDADLIEANLLWRQEMGGNITRVESENKVVILDSFBPLQAEBDEREVSVPA--EILRRS 2278

Qy  2133 RQFRARTGSLTLPPPPRSVPGVSCP-----------ESLQRSD----------PLEGPSNLPPSP 2176
Db  2279 RKF------PRAMPIWARPDYNPPLLESWKDPDYVPPVVHGCPLP-PAKAPPIP 2325

Qy  2177 P--------------------VLQLAMPMPLLGAGECNPF----TAIGCAMTETGGGPDDLP 2214
Db  2326 PPRRKRTVVLSESTVSSALAELA--TKTFGSSESSAVDSGTATASPDQPSDDGAGSDVE 2383

Qy  2215 SY---PP--KKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLG 2268
Db  2384 SYSSMPPLEGBPGDPDLSDGSWSTVSBEAS-------------------- 2413

Qy  2269 KSEFSCCSMSYTWTD-VISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTI 2327
Db  2414 EDVVCCSMSYTWTGALLITPCAAEETKLPINALSNSLLRHHNLVYATTSRSASLRQKKVTF 2473

Qy  2328 NRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAA 2387
Db  2474 DRLQVLDDHYRDVLKEMKAKASTVKAKLLSVEBEACKLTPPHSARSKF-GYGAKDVRN-LS 2531

Qy  2388 RKAVLDLQK----CVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRC 2443
Db  2532 SKAVNHIRSVWKDLLEDTETPI---DTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRV 2588

Qy  2444 VEKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMWSPD--AVGATCDTVCFDSTIT 2500
Db  2589 CEKMALYDVVSTLPQAVMGSSYGFQYSPGQRVEFLVNAWKAKKCPMGFAYDTRCFDSTVT 2648

Qy  2501 PEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTSS 2560
Db  2649 ENDIRVBESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKQQNCGYRRCRASGVLTTSC 2708

Qy  2561 SNSLTCWLKVNAAAEBQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAP 2620
Db  2709 GNTLTCYLKAAAACRAAKLQDCTMLVCGDDLVVICESAGTQEDEASLRAFTEAMTRYSAP 2768

Qy  2621 QDCVQPKYSLBELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWI 2680
Db  2769 PGDPPKPEYDLELITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPVNSWL 2828

Qy  2681 GYLIHHYPCLMVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGI 2740
```

```
Db   2829  GNIIMYAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGACVSIEPLDLPQIIQRLHGL  2888
Qy   2741  EAPSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKCARAVLASAKRRGGAHAKLARFLL-WH  2799
Db   2889  SAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWA  2948
Qy   2800  ATSR-----PLP------DLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFIVKYLAVI  2850
Db   2949  VRTKLKLTPIPAASQLDLSSWFVAGYS-----------GGDIYHSLSRARPRW-------F  2991
Qy   2851  VFALGLIAVGLAI  2863
Db   2992  MWCLLLLSVGVGI  3004
```

RESULT 12
US-09-539-601-33
; Sequence 33, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 3010
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-33

Query Match   21.2%; Score 3261; DB 4; Length 3010;
Best Local Similarity 29.8%; Pred. No. 2.8e-271;
Matches 968; Conservative 449; Mismatches 1191; Indels 642; Gaps 114;

```
Qy      8  TSPVPAPRTRKNKQTQASYPVSIK-----------TSVERGQRAKRKVQRDARPR        51
Db      3  TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR   59
Db     52  NYKIAGIHDGLQTLAQALP----------------AHGWGRQD                   79
Qy     60  GRR--------QPIPKARQPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPTD  111
Qy     80  PRHKSRNLGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG---  136
Db    112  PRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLP  170
Db    137  WFGVHLFVVCLLS-LACPCSGARVTDPTNTTLTNCCQRNQVIYCSPSTCLHEPGCVIC    195
Qy    171  GCSFSIFLLALLSCLTIPASAYEVRNV-SGVHVHTNDCSNASIVYEAADMIMHTPGCVPC  229
Qy    196  ADE-----CWPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVG  250
Db    230  VRENNSSRCWVALTPTLA-ARNASVPTTIRRHVDLLVGAAALCSAMYVGDLCGSVFIVA   288
Db    251  DMLV----RHWLIHIDLNETGTCYLEVPTGIDPGFL--GFIGW----MAGKVBAVIFLTKL  301
Qy    289  QLFTFSPRRHETVQ-DCN------CSIYPGHVTGHRMAWDMMMNWSPTAALVVSQL      337
Qy    302  ASQVPYAIATMFSSVHYLAVGALIYYASRGKWVQLLLALMLYIEATSGNPIRVPTGCSIA  361
Db    338  L-RIPQAVVDMVAGAHGVLAGLAYYSMVGNWAKVLIVMLLFAGVDGGTYV---TGGTMA   393
Db    362  E------FCSPLMIPCPCHSYLSE                                      379
Qy    394  KNTLGITLSFSPGSSQKIQLVNTNGSWHINRTALNCNDSLNTGFLAALFY---VHKFNSS  450
Qy    380  ------NVSEVICYSPKWTRPITLEYNNSISWYPY----TIPGARGC------         416
```

```
Qy    451  GCPERMASCSPIDAFAQGW-GPITYNESHSSDQRPYCWHYAPRPCGIVPAAQVCGPVYCF  509
Db    417  ------------MVKFKNNT-----WGCCRIRNVPSY------------C            437
Qy    510  TPSPVVVGTTDRFGVPTYSWGENETDVLLLNNTRPPQGNWFGCTWMNSTGFTKTCGGPPC  569
Db    438  TMG-------------TDAVWNDTRNTYEACGVTPWLT------------TAWHNGSALKLAILQ-  477
Qy    570  NIGGIGNKTLTCPFDCFRKHPEATYTKCGSGPWLTPRCLVHYPYRLWHYPCTVNFTIFKV  629
Db    478  --YPGSKE--MFKPHNWMSG-HLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVR  532
Qy    630  RMYVGGVEHRLEAACNWTRGERCNLEDRDR------SELSPLLLSTTEWQVLPCS------  678
Qy    533  GSWLQVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCG  592
Db    679  ------FTTLPALSTGLIHLHQNVDVQYLYGIGSAVVSFAIKWEYVLLLFLLLAD        728
Db    593  SKYLILAYLCYLSLCFGRASGYPLRPVLPSQSYILQAGWDVLSKAQVAPPALIFFICCYLR  652
Qy    729  ------IAQAEAALENLVV-------                                     753
Qy    653  CRLRYAALLGFVPMAAGLPLTFFVAAAAAQPDYDWVR-LLVAGLVLWAGRNGHRIALL    711
Db    754  --LNAASVAG---AHGI-LSFLVFFCAA----WYIKGRLVPGAAY------AL          790
Db    712  VGPWPLVALLTLLHLVTPASAFDTEI-IGGLTIPPVVALVVMSRFGFFAHLLPRCALV    768
Qy    791  YGVWPL-LLLLLALPPRAYAMDREMAASCGGAVFVGLILLTLSPHYKLFLARL------    842
Qy    769  NSYLWQRWENWFMNVTLRPERFFLVLVCFP----GATYDALVTFCVCHVALL-CLTSSA   822
Db    843  ---IW-------WLQYFITRAEAHLQVWI--PPLNVRGGRDAVILLTCAIHPELIFTITKIL  892
Qy    823  ASFFG------TDSRVRAHR------MLVRLGKCHAWYSHYVLKFFLLVFGENGVF       866
Db    893  LAILGPLMVLQAGITKVPYFVRAHGLIRACMLVR----KVAGGHYVQMALMKLAALTGTY  948
Qy    867  FYKHL------HGDVLPNDFASKLPLQEPFPFPEGKARVYRNEGRRLACGDTVDGLPVV   919
Db    949  VYDHLTPLRDWAHAGL--RDLAVAV---EPVVFSDMETKVITWGADTAACGDILLGLPVS 1003
Qy    920  ARLGDLVFAGLA--MPPDGWATTAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLG  977
Db   1004  ARRGREIHLGPADSLEGCGWRLLAPITAYSQQTRGLLGCIITSLTGRDRNQVEGEVQVVS 1063
Qy    978  SLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN-DQDI--YQPPCGAGSL 1034
Db   1064  TATQSFLATCVNGVCWTVYHGAGSKTLLAGPKG---PITQMYTNVDQDLVGWQAPPGARSL 1120
Qy   1035  TRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMF 1094
Db   1121  TPCTCGSSDLYLVTRHADVIPVRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIF  1180
Qy   1095  TAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILLAPTGSGKSTKL 1154
Db   1181  RAAVCTRGVAKAVDFVPVESMGTTMRSPVFTDNSSPPAVPQTFQVAHLHAPTGSGKSTKV 1240
Qy   1155  PLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYL 1214
Db   1241  PAAVAGQGVKVLVLNPSVAATLGFGAYMSKAHGIDPNIRIGVRTITTGAPITYSTYGKFL 1300
Qy   1215  T-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLLVLATATPPGVIPTPHA 1272
Db   1301  ADGGCSGGAYDIIICDECHSTDSTTILGIGTVLDQAETAGARLVVLATATPPGSVTVPHP 1360
Qy   1273  NITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELARKGITAVSYYRG 1332
Db   1361  NIEVALSSTGEIPFYGKAIPIETIKGGRHLIFCHSKKKCDELAAKLSGLGLNAVAYYRG  1420
Qy   1333  CDISKIP-EGDCVVVATDALCTGYTGDFDSVVDCLSMVEGTCHVDLDPTFTMGVRVCGVS 1391
Db   1421  LDVSVIPTSGDVIVATDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTVPQD   1480
```

```
Qy  2331 PLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKA 2390
          |||      ::  |||:  :      ::  ||    |:  =||
Db  2477 QVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPHSARSKF-GYGAKDVRN-LSSKA 2534

Qy  2391 VLDLQK----CVEAGEIPSHYRQTVIVPKBEVFVKTQKPTKKPPRLISYPHLEMRCVEK 2446
          :  ::       |||   :     : ||            ||   :|| |||  ||||:
Db  2535 VNHIRSVWKDLLEDTETPI--DTTIMAKNEVFCVQPEKGGRKPARLLVFPDLGVRVCEK 2591

Qy  2447 MYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMWSPD--AVGATCDTVCFDSTITPED 2503
          |  ||  ||  ||||| :|||   |  ||||| || |      ||| || |||||:| |
Db  2592 MALYDVVSTLPQAVMGSSYGFQYSPSGQRVEFLVNAWKAKKCPMGFAYDTRCFDSTVTEND 2651

Qy  2504 IMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVVTTSSSNS 2563
          | ||:|||  :       |  : |  ||| ||| :  :|  ||||||:||| |||  :|:
Db  2652 IRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNT 2711

Qy  2564 LTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDC 2623
          |||:||   |  |  :|     :||||| |   |||  :||  : ||: | |  |||  |
Db  2712 LTCYLKAAAACRAAKLQDCTMLVCGDDLVICESAGTQEDEASLRAFTEAMTRYSAPPGD 2771

Qy  2624 VPQPKYSLEELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYL 2683
          :|:|:| ||:|:|||||| : ||||||  ||||||  ||| |::    ||||| |||| :
Db  2772 PPKPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNI 2831

Qy  2684 IHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAF 2743
          |  | ||| ||||| ||::: | : :  | : :  ||:|   ||||| ||| |||| |||
Db  2832 IMYAPTLWARMLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGLSAF 2891

Qy  2744 SVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WHATS 2802
          |   |:  |: ||:  |  ||:|||| || :||:  : ||||| |  | | :|  | |
Db  2892 SLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVRT 2951

Qy  2803 R----PLP----DLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLVAVIVFA 2853
          |     ||    ||| || : |  ::||| |||| |
Db  2952 KLKLTPIPAASQLDLSSWFVAGYS-----------GGDIYHSLSRARPRW------FWMC 2994

Qy  2854 LGLIAVGLAI 2863
          | |:|:|: |
Db  2995 LLLLSVGVGI 3004


RESULT 13
US-09-539-601-27
; Sequence 27, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 3010
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-27

Query Match         21.2%;  Score 3255;  DB 4;  Length 3010;
Best Local Similarity   29.8%;  Pred. No. 9.3e-271;
Matches  968;  Conservative  453;  Mismatches 1184;  Indels 648;  Gaps 116;

Qy   8 TSPVPAPTRKNKQTQASYPVSIK----------TSVERGQRAKRKVQRDARPR 51
         |:| |: |::|                       :|::|||  |  ||:|: :
Db   3 TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR 59

Qy  52 NYKIAGIHDGLQTLAQAALP------------------AHGWGRQD 79
         :|::    | |                          | |  |:
Db  60 GRR-------QPIPKARQPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPTD 111
```

```
Qy  1392 AIVKGQRRGRTGRGRAGIYYVYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTILD 1451
          | :  |||||||||| |||:   ||| |||| :  :|  | |||| ||  |:   | ||
Db  1481 AVSRSQRRGRTGRGRMGIYRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAETSVRLR 1540

Qy  1452 TYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLQLCHQYGYA 1510
           | | ||||  :  | ||| |   :|  |||| || || ||:||   ||: |   :|
Db  1541 AYLNTPGLPVCQDHLEFWESVFTGLTHIDAHFLSQTKQAGDNFPYLVAYQATVCARAQAP 1600

Qy  1511 APNDAPRWQG-ARLGKK---PCGVLWRLDGADACPGPEPSEVTRYQM-CFT-----EVNTSG 1562
          :|  : |||  ||:|    | | |  |  || :  ||  :: ||    |:        ||
Db  1601 PPSWDQMWKCLIRLKPTLHGFTPLLYRL-GAVQNEVTTTHPITKYIMACMSADLEVVTS- 1658

Qy  1563 TAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI----VEECA 1616
           : | || | : ||   :    |    | ||   |  |:    || :    ||||
Db  1659 TWVLVGGVLAALAAYCLTTGSSVIVGR------IILSGKPAIIPDREVLYREFDEMEECA 1712

Qy  1617 SFIP-LEAMVAAIDKLK-------STIT----TTSPFTLETALEKLNTFLGPHAAATILAI 1664
          |  | :|        :         :|      || || |  ||| ||| |  ||:
Db  1713 SHLPYIEQGMQLAEQFKQKAIGLLQPATKQAEAAAP-VVESKWRTIEAFWAKHMWNNFISG 1771

Qy  1665 IEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGATASKLTDARGALAFMM 1724
          | | |||:||| ||| :  :||  :||:||   |  :  |  | :||:|| ||| ||  |
Db  1772 IQYLAGLSTLPGNPAIASLMAFTASITSPLTTQHTLLFNILGGWVAAQLAPPSAASAFVG 1831

Qy  1725 AGAAGTALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVVSAFN 1781
          || || | | |  |    :   :|||  ||  : :   | :|||  |::   | ::| |
Db  1832 AGIAGAAVG---SIGLGKVLVDILAGYGAGVAGALVAFKVMSGELPSTEDLVNLLPAILS 1888

Qy  1782 PAAGVGVLSACAMPALTTAGPD---HWPNRLLTMLARSNTVCNEYFIATRDIRRKILG 1837
          | | |||:  || :|:|  |     :| ||| ::| || : |::   :  |   :
Db  1889 PGALVVGVV--CAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAARVTQ 1946

Qy  1838 ILEASTPWSVISACIRWLHTPTEDDCGLI----AWGLEIWQYVCNFFVICFNVLKAGVQS- 1893
          || :  ||:|:|| |||| | ::| |       ||: : |  ||:|| || :| |: |
Db  1947 ILSSLT----ITQLLKRLHQWINEDCSTPCSGSWLRDVWDWICT----VLTDFKTWLQSK 1998

Qy  1894 -MVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWR 1952
           : :|||||::||::|||| | | |:||  ||||| :|||||||| ||:|||||||:| :
Db  1999 LLPRLPGVPFFSCQRGYKGVWRGDGIMQTTCPCGAQITGHVKQNGSMRIV-GPRTCSNTWH 2057

Qy  1953 GAVPVNARLCGSARPDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFTQVP 2010
          | .|:||   |: | |   ::  |: :  | |:  ||  ||:| ||  :::|   ||| |
Db  2058 GTFPINAYTTGPCTPSPAPNYSRALMRVGAEEYVEVTRVGDFHYVTGMTTDNVKCPCQVP 2117

Qy  2011 -------------PTLRAAVA---------VDGVQVQCYLGEPKTPWTTSACC 2041
                        || ||          ||| ||| |||  ||| ||:
Db  2118 APEFFTEVDGVRLHRYAPACKPLLREEVTFLVGLNQYLVGSQLPC---EPE------- 2165

Qy  2042 YGPDGKGKTVKL--PPRVDGHTPGVRMQLNLRDALETND----------CNSTNNTPS 2087
            ||      ||  || : |   |  : | ||| ::::           |||:   |
Db  2166 --PDVAVLTSMLTDPSHITAETAKRRLARGSPPPLASSSASQLSAPSLKATCTTRHDSPD 2223

Qy  2088 DEAAVSALVFKQEL--------RRTNQLL-----EAISAGVDTTKLPAPSIEEVVRKRQF 2135
          |      |  ||          ||:|          |:|| || |||||| ||: :|:|
Db  2224 ADLIEANLLWRQEMGGNITRVESENKVVILDSFEPLQAEEDERGVSVPA--EILRRSRKF 2281

Qy  2136 RARTGSLTLPPPPRSVPGVSCP--------ESLQRSD--------PLEGPSNLPPSPP-- 2177
            || ||| ||||| ||| :||        |||  |          ||||| :||||
Db  2282 PRAMPIWARPDYNPPLLESWKDPDYVPPVVHGCPLP-PAKAPPIPPSR 2328

Qy  2178 -------------VLQLAMPMPLLGAGECNPF-----TAIGCAMTETGGGPDDLPSY- 2216
                       | | |  | |:| || :|:     ||:| : | :  |  || ||
Db  2329 RKRTVVLSESTVSSALAELA--TETFGSSESSAVDSGTATASPDQPSDDGDAGSDVESYS 2386

Qy  2217 --PP---KKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSE 2271
            ||   |:| |||:  |||  :  :|:|  :|   |   :|    ::
Db  2387 SMPPLEGEPGDPDLSDGSWSTVSEEAS---------EDV 2416

Qy  2272 FSCSMSYTWTD-VISFKTASKVLSATRAITSGFLKQRSLVVTEPRDAELRKQKVTINRQ 2330
          : ||||||||  :|      ||| || |||| ||::|:| :| ||:| ||||||| ||
Db  2417 VCCSMSYTWTGALIITPCAAEETKLPINALSNSLLRHHNLVYATTSRSANLRQKKVTFDRL 2476
```

```
Qy    80  PRHKSRNLGILLDYPLGWIGDVTHTPLVGAVVRPVCQIVRLLEDGVNWATG---        136
Db   112  PRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLP  170

Qy   137  WFGVHLFVVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC  195
Db   171  GCSFSIFILLALLSCLTIPASAYEVRNV-SGVVHVTNDCSNASIVYEAADMIMHTPGCVPC  229

Qy   196  ADE------CWVPANPYISHPSNWTGTDSFLADHIDFWMGALVTCDALDIGELCGACVLVG  250
Db   230  VRENNSSRCWVALTPTLA-ARNASVPTTIRRHVDLLVGAAALCSAMYVGDLCGSVFLVA   288

Qy   251  DWLV-----RHWLIHIDLNETGTCYLEVPTGIDPGFL--GFIGW---MAGKVBAVIFLTKL  301
Db   289  QLFTFSPRRHETVQ-DCN-----------CSIYPGHVTGHRMAWDMMMNMNWSPTAALVVSQL  337

Qy   302  ASQVPYAIATMPSSVHYLAVGALIYASRGKNYQLLLALMLYIEATSGNPIRVPTGCSIA   361
Db   338  L-RIPQAVDMVAGAHHNGVLAGLAYYSMVGNWAKVLIVMLLFAGVDGGTYV---TGGTMA  393

Qy   362  E-----------------------------------FCSPLMIPCPCHSYLSE         379
Db   394  KNTLGITSLFSPGSSQKIQLVNTNGSWHINRTALNCNDSLNTGFLAALFY---VHKFNSS  450

Qy   380  -----------NVSEVICYSPKWTRPITLEYNNSISWYPY-----------TIPGARGC-  416
Db   451  GCPERMASCSPIDAFAQGW-GPITYNESHSSSDQRPYCWHYAPRPCGIVPAAQVCGPVYCF  509

Qy   417  ------------------MVKFKNNT----------WGCCRIRNVPSY--------C     437
Db   510  TPSPVVVGTTDRPGVPTYSWGENETDVLLLNNTRPPQGNWFGCTWMNSTGFTKTCGGPPC  569

Qy   438  TMG-----------TDAVWNDTRNTYEACGVTPWLT-----------TAWHNGSALKLAILQ-  477
Db   570  NIGGIGNKTLTCPTDCPRKHPEATYTKCGSGPWLTPRCLVHYPYRLMHYPCTVNFTIFKV  629

Qy   478  --YPGSKE--MFKPHNWMSG-HLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVR  532
Db   630  RMYVGGVEHRLEAACNWTRGERCNLEDRDR------SELSPLLLSTTEWQVLPCS------  678

Qy   533  GSWLQVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCG  592
Db   679  ----------FTTLPALSTGLIHLHQNVVDVQYLYIGISAVVSFAIKWEYVLLLFLLLAD  728

Qy   593  SKYLILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLR  652
Db   729  ARVCACLWMMLL----------------------IAQAEAALENLVV------------   753

Qy   653  CRLRYAALLGFVPMAAGLPLTFFVAAAAQPDYDWWVR--LLVAGLVLWAGRNRGHRIALL  711
Db   754  --LNAASVAG------AHGI-LSFLVFFCAA-------WYIKGRLVPGAAY--------AL  790

Qy   712  VGPWPLVALLTLHLVTPASAFDTEI---IGGLTIPPVVALVVMSRFGFFAHLLPRCALV   768
Db   791  YGVWPL--LLLLLALPPRAYAMDREMAASCGGAVFVGULLLTILSPHYKLFLARL------  842

Qy   769  NSYLWQRWENWFWNVTLRPERRFLVLVCFP-----GATYDALVTFCVCHVALL-CLTSSA  822
Db   843  ---IW-----WLQYFITRAEAHLQVWI--PPLNVRGGRDAVILLTCAIHPELIFTITKIL  892

Qy   823  ASFFG----------TDSRVRAHR------MLVRLGKCHAWYSHYVLKFFLLVFGENGVF  866
Db   893  LAILGPLMVLQAGITKVPYFVRAHGLIRACMLVR-----KVAGGHYVQALMKLAALTGTY  948

Qy   867  FYKHL--------HGDVLPNDFASKLPLQEPFFPFEGKARVYRNEGRRLACGDTVDGLPVV  919
Db   949  VYDHLITPLRDWAHAGL--RDLAVAV--EPVVFSDMETKVITWGADTAACGDIILGLPVS 1003

Qy   920  ARLGDLVFAGLA--MPPDGWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLG  977
Db  1004  ARRGREIHLGPADSLEGQGWRLLAPITAYSQQTRGLLGCIITTSLTGRDRNQVEGEVQVVS 1063

Qy   978  SLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN-DQDI--YQPPCGAGSL 1034
```

```
Db  1064  TATQSFLATCVNGVCWTVYHGAGSKTLAGPKG---PITQMYTNVDQDLVGWQAPPGARSL 1120
Qy  1035  TRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMF 1094

Db  1121  TPCTCGSSDLYLVTRHADVIPVRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIF  1180
Qy  1095  TAA---RNSGGGVSQIRVRPLVCAGYHPQYTAHATLDKTPTVNEYSVQILIAPTGSGKS  1151

Db  1181  RAAVCTRGVAKAVDFVPVESMETTMRSPVFTDNS---SPPAVPQTFQVAHLHAPTGSGKS 1237
Qy  1152  TKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYPNGKCNTNGASLTYSTYG 1211

Db  1238  TKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRIGVRTITTGAPITYSTYG 1297
Qy  1212  MYLT-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPT 1269

Db  1298  KFLADGGCSGGAYDIIICDECHSTDSTTILGIGTVLDQAETAGARLVVLATATPPGSVTV 1357
Qy  1270  PHANITEIQLTDEGTIPFHGKVIKEENLKKGRHLIFEATKKHCDELANELARKGITAVSY 1329

Db  1358  PHPNIEEVALSSTGEIPFYGKAIPIETIKGGRHLIFCHSKKKCDELAAKLSGLGLNAVAY 1417
Qy  1330  YRGCDISKIP-EGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVC 1388

Db  1418  YRGLDVSVIPTSGDVIVVATDALMTGFTGDFDSVIDCNTCVTQVDFSLDPTFTIETTTV  1477
Qy  1389  GVSAIVKGQRRGRTGRGRAGIYYVVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQT 1448

Db  1478  PQDAVSRSQRRGRTGRGRMGIYRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAETSV 1537
Qy  1449  ILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLQLCHQY 1507

Db  1538  RLRAYLNTPGLPVCQDHLEFWESSVFTGLTHIDAHFLSQTKQAGDNFPYIVAYQATVCARA 1597
Qy  1508  GYAAPNDAPRWQG-ARLGKK---PCGVLWRLDGADACPGPEPSEVTRYQM-CFT---EVN 1559

Db  1598  QAPPPSWDQMWKCLIRLKPTLHGPTPLLYRL-GAVQNEVTTHPITKYIMACMSADLEVV  1656
Qy  1560  TSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI----VE    1613

Db  1657  TS-TWVLVGGVLAALAAYCLTTGSVVIVGR------IILSGKPAIIPDREVLVREFDEME 1709
Qy  1614  ECASFIP-LEAMVAAIDKLK------STIT----TTSPFTLETALEKLNTFLGPHAATI  1661

Db  1710  ECASHLPYIEQGMQLAEQFKQKAIGLLQTATKQAEAAAP-VVESKWRTIEAFWAKHMWNF 1768
Qy  1662  LAIIEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIRMFLSLFGGAIASKLTDARGALA 1721

Db  1769  ISGIQYLAGLSLTLPGNPAIASLMAFTASITSPLTTQHTLLFNIIGGWVAAQLAPPSAASA 1828
Qy  1722  FMMAGAAGTALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYS 1778

Db  1829  FVGAGIAGAAVG---SIGLGKVLVDILAGYGAGVAGALVAFKVMSGEMPSTEDLVNLLPA 1885
Qy  1779  AFNPAAGVVGVLSACAMFALTTAGPD----HWPNRLLTMLARSNTVCNEYFIATRDIRRK 1834

Db  1886  ILSPGALVVGVV--CAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAAR 1943
Qy  1835  ILGILEASTPWSVISACIRWLHPTPEDDCGLI----AWGLEIWQYVCNPFVICFNVLKAGV 1891

Db  1944  VTQILSSLT----ITQLLKRLHQWINEDCSTPCSGSWLRDVWDWICT----VLTDFKTWL 1995
Qy  1892  QS--MVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSN 1949

Db  1996  QSKLLPRLPGVPFFSCQRGYKGVWRGDGIMQTTCPCGAQITGHVKNGSMRIV-GPRTCSN 2054
Qy  1950  YWRGAVPVNARLCGSARPDPT-DWTSLVVNYGVRDYCYKEKMGDHIFVTAVSSPNV-CFT 2007

Db  2055  TWHGTFPINAYTTGPCTPSPAPNVSRALWRVAAEEYVEVTRVGDFHYVTGMTTDDVKCPC 2114
Qy  2008  QVP-----------------------PTLRAAVA------------VDGVQVCYLGEPKTPWTTS 2038
```

```
;  TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
;  FILE REFERENCE: all sequences
;  CURRENT APPLICATION NUMBER: US/09/539,601C
;  CURRENT FILING DATE: 2001-08-30
;  EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
;  EARLIER FILING DATE: 1999-04-03
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: PatentIn Ver. 2.1
;  SEQ ID NO 21
;  LENGTH: 3010
;  TYPE: PRT
;  ORGANISM: Hepatitis C virus
US-09-539-601-21

Query Match      21.1%;   Score 3252;   DB 4;   Length 3010;
Best Local Similarity  29.9%;   Pred. No. 1.7e-270;
Matches  968;  Conservative 453;  Mismatches 1197;  Indels 622;  Gaps 115;

Qy    8  TSPVPAPRTRKNKQTQASYPVSIK----------------------------TSVERGQRAKRKVQRDARPR  51
Db    3  TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR  59

Qy   52  NYKIAGIHDGLQTLAQAALP--------------------------AHGWGRQD  79
Db   60  GRR--------QPIPKARQPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPTD  111

Qy   80  PRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG---  136
Db  112  PRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLP  170

Qy  137  WFGVHLFVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC  195
Db  171  GCSFSIFLALLSCLTIPASAYEVRNV-SGVYHVTNDCSNASTVYEAADMIMTPGCVPC  229

Qy  196  ADE-----CWYPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVG  250
Db  230  VRENNSSRCWVALTPTLA-ARNASVPTTTIRRHVDLLVGAAALCSAMYVGDLCGSVFLVA  288

Qy  251  DWLIV----RHWLIHIDLNETGTCYLEVPTGIDPGFL---GFIGW---MAGKVEAVIFLTKL  301
Db  289  QLFTFSPRRHETVQ-DCN----CSIYPGHVTGHRMAWDMMNWSPTAALVVSQL  337

Qy  302  ASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSIA  361
Db  338  L-RIPQAVVDMVAGAHWGVLAGLAYYSMVGNWAKVLIVMLLFAGVDGGTYV----TGGTMA  393

Qy  362  E--------------------------------------FCSPLMIPCPCHSYLSE  379
Db  394  KNTLGITSLFSPGSSQKIQLVNTNGSWHINRTALNCNDSLNTGFLAALFY---VHKFNSS  450

Qy  380  -------NVSEVICYSPKWTRPITLEYNNSISWYPY--------------------TIPGARGC  416
Db  451  GCPERMASCSPIDAFAGGW-GPITYNESHSSDQRPYCWHYAPRPCGIVPAAQVCGPVVCF  509

Qy  417  ----------------MVKFKNNT------------------WGCCRIRNVPSY-------C  437
Db  510  TPSPVVGTTDRFGVPTYSWGGNETDVLLLNNTRPPQGNWFGCTWMNSTGFTKTCGGPPC  569

Qy  438  TMG----------TDAVWNDTRNTYEACGVTPWLT--------TAWHNGSALKLAILQ-  477
Db  570  NIGGIGNKTLTCPTDCPRKHPEATYTKCGSGPWLTPRCLVHPYRLWHYPCTVNFTIFKV  629

Qy  478  --YPGSKE--MFKPHNWMSG-HLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVR  532
Db  630  RMYVGGVEHRLEAACNWTRGERCNLEDRDR------SELSPLLLSTTEWQVLPCS-----  678

Qy  533  GSWLQVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCG  592
Db  679  --------FTTLPALSTGLIHLHQNVVDVQYLYGIGSAVVSFAIKWEYVLLFLLLAD  728

Qy  593  SKYLILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYLR  652
Db  729  IAQAEAALENLVV------------ARVCACLWMMLL  753


Db 2115  QVPAPEFFTEVDGVRLHRYAPACKPLLREEVTFLVGLNQYLVGSQLPC---EPE------  2165
Qy 2039  ACCYGPDGKGKTVKL-PFRVDGHTPGVRMQLNLRDALETND----------  2084

Db 2166  -----PDVAVLTSMLTDPSHITAETAKRRLARGSPPPLASSSASQLSAPSLKATCTTRHD  2220
Qy 2085  TPSDEAAVSALVFKQEL--------RRTNQLL-------EAISAGVDTTKLPAPSIEEVVVRK  2132

Db 2221  SPDADLJEANLLWRQEMGGNITRVESENKVVILDSFEPLQAEBEDEREVSVPA--EILRRS  2278
Qy 2133  RQFRARTGSLTLPPPPRSVPGVSCP---------PLEGPSNLPPSP  2176

Db 2279  RKF------------PRAMPIWARPDYNPLLESWKDPDYVPPVVHGCPLP-PAKAPPIP  2325
Qy 2177  P------------VLQLAMPMPLLGAGECNPF-----TAIGCAMTETGGGPDDLP  2214

Db 2326  PSRRRKRTVVLSESTVSSALAELA--TETFGSSESSAVDSGTATASPDQPSDDGDAGSDVE  2383
Qy 2215  SY---PP---KKEVSEWSDESWSTATTASSVVTGPPYPKIRGKDSTQSAPAKRPTKKKLG  2268

Db 2384  SYSSMPPLEGEPGDPDLSDGSWSTVSEEAS-----------  2413
Qy 2269  KSEFSCCSMSYTWTD-VISFKTASKVLSATRAITSGFLKQRSLVYYTEPRDAELRKQKVTI  2327

Db 2414  EDVCCSMSYTWTGALITPCAAEETKLPINALSNSLLRHHNLVYATTSRSASLRQKKVTF  2473
Qy 2328  NRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAA  2387

Db 2474  DRLQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPHSARSKF-GYGAKDVRN-LS  2531
Qy 2388  RKAVLDLQK----CVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRC  2443

Db 2532  SKAVNHIRSVWKDLLEDTETPI---DTTIMAKNEVFPCVQPEKGGRKPARLIVFPDLGVRV  2588
Qy 2444  VEKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMWSPD--AVGATCDTVCFDSTIT  2500

Db 2589  CEKMALYDVVSTLPQAVMGSSYGFQYSPGQRVEFLVNAWKAKKCPMGFAYDTRCFDSTVT  2648
Qy 2501  PEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGMIAYDGREIGYRRCRSSGVYTTSS  2560

Db 2649  ENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSC  2708
Qy 2561  SNSLTCWLKVNAAAEQAGMKNPRFLICGGDCTVIWKSAGADADKQAMRVPASWMKVMGAP  2620

Db 2709  GNTLTCYLKAAAACRAAKLQDCTMLVCGDDLVVICESAGTQEDEASLRAFTEAMTRYSAP  2768
Qy 2621  QDCVPQPKYSLEELTSCCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWI  2680

Db 2769  PGDPPKPEYDLELITSCSSNVSVAHDASGKRVYYLITRDPTTPLARAAWETARHTPVNSWL  2828
Qy 2681  GYLIHYPCLWVSRVLAVHFMEQMLPEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGI  2740

Db 2829  GNIIMYAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGL  2888
Qy 2741  EAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WH  2799

Db 2889  SAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWA  2948
Qy 2800  ATSR-----PLP----DLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVI  2850

Db 2949  VRTKLKLTPIPAASQLDLSSWFVAGYS---------GGDIYHSLSRARPRW------F  2991
Qy 2851  VFALGLIAVGLAI  2863

Db 2992  MWCLLLLSVGVGI  3004
```

```
RESULT 14
US-09-539-601-21
; Sequence 21, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
```

```
Qy 1662 LAIIEYCCGLVTLPDNPPASCVPAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALA 1721
        ::.| ||  ::.|:  ||  ||:|..|:.: .  :|.|.:::|||.|:  . .:.|.|:
Db 1769 ISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTTQHTLLFNILGGWVAAQLAPPSAASA 1828

Qy 1722 FMMAGAAGTALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYS 1778
        |.  |.||.| |   ..||    |  |.||.|| |:|| : ||:.||: . :::. ::
Db 1829 FVGAGIAGAAVG--SIGLGKVLVDILAGYGAGVAGALLVAFKVMSGEMPSTEDLVNLLPA 1885

Qy 1779 AFNPAAGVVGVLSACAMFALTTAGPD----HWPNRLLTMLARSNTVCNEYFIATRDIRRK 1834
         :.|.||:|||:  ||::   . |     .:.||||:::|:| |.   ::..:: :||
Db 1886 ILSPGALVVGVV--CAAILRRHVGPGEGAVGQWMNRLIAFASRGNHVSPTHYVSESDAAAR 1943

Qy 1835 ILGILEASTPWSVISACIRWLHTPTEDDCGLI----AWGLEIWQVCNFPVICFNVLKAGV 1891
        ::.||  |   .:.:|| ||:|: ::| |.:     :|.||:||:| .| :|.:.|
Db 1944 VTQILSSLT----ITQLLKRLHQWINEDCSTPCSGSWLRDVWDWICT---VLTDFKTWL 1995

Qy 1892 QS--MVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSN 1949
        ||   : ||.|||||||:||||.|.| |::|: ||||| :| .::||.:::..|||||||
Db 1996 QSKLLPRILPGVPFFSCQRGYKGVWRGDGIMQTTCPCGAQITGHVKNGSMRIV-GPRTCSN 2054

Qy 1950 YWRGAVPVNARLCGSARPDT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFT 2007
        .| |. |:||   |: .|   .. |:..|||.|.|.|::::||||:||::  .:| |.|
Db 2055 TWHGTFPINAYTTGPCTPSPAFNYSRALWRVAAEEYVEVTRVGDFHYVTGMTTDNVKCPC 2114

Qy 2008 QVPPTLRAAVAVDGVQVQCY------LGEPKT-----------PWTTSACCYGPDGKGKTV 2051
        ||| ..  .|.:|||::: |       .||.:            | |:|||.|||||::|
Db 2115 QVPAP-EFFTEVDGVRLHRYAPACKPLLREEVTFLVGLNQYLVGSQLPCGPEPDVAVLTS 2173

Qy 2052 KL--PFRVDGHTPGVRMQLNLRDALETND----------CNSTNNTPSDEAAVSALVF 2097
        :|    ..|.|: .:|::|.:| |||..           |.|  : |.||.::|:||
Db 2174 MLTDPSHITAETAKRRLARGSPPSLASSSASQLSAPSLKATCTTRHDSPDADLIEANLLW 2233

Qy 2098 KQEL-------EAISAGVDTKLPAPSIEEVVRKRQFRARTGSLTLP 2145
        :||.       | |:. |.|:.  |.:| ::  : .|.|:   :::
Db 2234 RQEMGGNITRVESENKVVILDSFEPLQAEEDEREVSVPA--EILRRSRKF------- 2281

Qy 2146 PPPRSVPGVSCP---------ESLQRSD----------PLEGPSNLPPSPP------ 2177
        |  |::|..  |         ||::.:           |   |..||| ||
Db 2282 --PRAMPIWARDPDYNPPLLESWKDPDYVPPVHGCPLP-PAKAPPIPPPRRERTVVLSES 2338

Qy 2178 ------VLQLAMPMPLLGAGECNPF-----TAIGCAMTETGGGPDDLPSY---PP---KKE 2221
              |:.:|  .. .||| |.        || |..: :|||.| :||..        ::|
Db 2339 TVSSALAELA--TKTFGSSESSAVDSGTATASPDQPSDDGDAGSDVESYSSMPPLEGEPG 2396

Qy 2222 VSEWSDESWSTATTASSVVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTWT 2281
        .:.|| ||||:  :.          | :.|:.           ::    | |.|||||||
Db 2397 DPDLSDGGSWSTVSEEAS--------------------EDVVCCSMSYTWT 2426

Qy 2282 D-VISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ 2340
        . |:|:.::|::|.| |:.:.:.|:: :||:||  |   .|::||| :| |  |  .: :.
Db 2427 GALITPCAAEETKLPVNALSNSLLRHHNLVYATTSRSASLRQKKVTFDRLQVLDDHYRDV 2486

Qy 2341 VRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQK---- 2396
        ::  :|||:.|  :  . :|.  .||::||||:::||||||:| ||  :|:::::
Db 2487 LKEMKAKASTVKAKLLSVEEACKLTPPHSARSKF-GYGAKDVRN-LSSKAVNHIRSVWKD 2544

Qy 2397 CVEAGEIPSHYRQTVIVPKEEVVKTPQKPTKKPPRLLSYPHLEMRCVEKMYGQVAPDV 2456
        |  |.||.|::.:|.:. ||:|.:.||:  :|| ||:..|.| .||||||.  |...:
Db 2545 LLEDTETPI---DTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVSTL 2601

Qy 2457 VKAVMGDAYGF-VDPRTRVKRLLSMWSPD--AVGATCDTVCFDSTITPEDIMVETDIYSA 2513
        . |||| :||| .  :.||.|.:..|    |:|.||||.|||||.:| |||.||.|||.
Db 2602 PQAVMGSSYGFQYSPGQRVEFLVNAWKAKKCPMGFAYDTRCFDSTVTENDIRVEESIYQC 2661

Qy 2514 AKLSDQHRAGIHTIARQLYAGGMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAA 2573
        .:|:||:::||:.:.|.|||||::::.|::  |||||:||| ||| .|:|||:||..||
Db 2662 CDLAPEARQAIRSLFERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKAAAA 2721

Qy 2574 AEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEE 2633
        . :|:::|.|||:||||  |||:||| .|::||:|.|..|||.||||.||.|:|:|.||
Db 2722 CRAAKLQDCTMLVCGDDLVVICESAGTQEDEASLRAFTEAMTRYSAPGDPPKPEYDLEL 2781

Qy 2634 LTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVS 2693
```

```
Qy 653 CRLRYAALLGFVPMAAGLPLTFFVAAAAAQPPDYDWVR-LLIVAGLVLWAGRNRGHRIALL 711
         ::|  |: || | :| : :.|:. ||         :   |::||||:|||:||||:|||
Db 754 --LNAASVAG---AHGI-LSFLVFFCAA-----WYIKGRLVPGAAY------AL 790

Qy 712 VGPWPLVALLTLLHLVTPASAFDTEI-IGGLTIPPVVALVMSRFGFFAHLLPRCALV 768
        .|.|||  ||.|| :: |.:|.  .: .||: :. :: ||: ::  .|||:.||
Db 791 YGVWPL-LLLLALPPRAYAMDREMAASCGGAVFVGLILLTLSPHYKLFLARL------ 842

Qy 769 NSYLWQRWENWFWNVTLRPERFFLVLCFP-----GATYDALVTFCVCHVALL-CLTSSA 822
        :  :|   .:.. :: :   .:.::||.:      |.| |:| .|.|||:|: ||:::
Db 843 --IW----WLQYFITRAEAHLQVWI--PPLNVRGRGRDAVILLTCAIHPELIFTITKIL 892

Qy 823 ASFFG-------TDSRVRAHR-----MLVRLGKCHAWYSHYVLKFFLLVFGENGVF 866
        |:.:|       |.|  ||       ||.||.. .|| :||||:|.|::.| :.
Db 893 LAILGPLMVLQAGITKVPYFVRAHGLIRACMLVR----KVAGGHYVQMALMKIAALTGTY 948

Qy 867 FYKHL------HGDVLPNDFASKLPLQEPFPFEGKARVYRNEGRRLACGDTVDGLPVV 919
        .| ||      .|| |. .  :..  .|...| ||.::   :..:| |||.:.|||||
Db 949 VYDHLTPLRDWAHAGL--RDLAVAV---EPVVFSDMETKVITWGADTAACGDIILGLPVS 1003

Qy 920 ARLGDLVFAGLA--MPPDGWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLG 977
        ||.|.:. .| .  ..|:.  : ||.| ..|:||:|. :.. :|::|| |.:.||:..:|
Db 1004 ARRGREIHLGPADSLEGQGWRLLAPITAYSQQTRGLLGCIITSLTGRDRNQVEGEVQVVS 1063

Qy 978 SLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN-DQDI--YQPPCGAGSL 1034
        : |.|::.|||:.| .||:||::. .|| |.|.: |:|..:.: |||:  : .|.||.||
Db 1064 TATQSFLATCVNGVCWTVYHGAGSKTLAGPKG---PITQMYTNVDQDIVGWQAPPGARSL 1120

Qy 1035 TRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMF 1094
        | |:||.:. |||||.. ::.. |.|| || ||.||:.::|||||||:||.|||:||:|
Db 1121 TPCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIF 1180

Qy 1095 TAA---RNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILLAPTGSGKS 1151
         ||   ||:|. |.:.  .|||.|.  ..:.:  :.: |.   ::: |:|::||||||||
Db 1181 RAAVCTRGVAKAVDFVPVESMETTMRSPVFTDNS---SPPAVPQTFQVAHLHAPTGSGKS 1237

Qy 1152 TKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYG 1211
        ||.|::|.:: |:|||||||||:|.:. .||:|::||:|: | .|.. :||| :|||||
Db 1238 TKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVGTITTGAPITYSTYG 1297

Qy 1212 MYLT-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPT 1269
          |  | || . ||:|||||| :|::|::||||.||:.| ::.||:|||||||||.:.
Db 1298 KFLADGGCSGGAYDIIICDECHSTDSTTILGIGTVLDQAETAGARLVVLATATPPGSVTV 1357

Qy 1270 PHANITEIQLTDEGTIPPHGKKIKEENLKKGRHLIFEATKKHCDELANELARKGITAVSY 1329
        || |:.|: ::..|.||::|| | |.||:.||||||.::|| |||||.:|: :|: |:|
Db 1358 PHPNIEEVALSSTGEIPFYGKAIPIATIKGGRHLIFCHSKKKCDELAAKLSGLGLNAVAY 1417

Qy 1330 YRGCDISKIP-EGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVC 1388
        |||.|:|.|  .||.:|||||||.||:||||||| ||.:.  |.|.:|||||| . :.
Db 1418 YRGLDVSVIPTSGDVIVVATDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTV 1477

Qy 1389 GVSAIVKGVQRRGRTGRGRAGIYYVVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQT 1448
         : |:::|.|||||||||| ||| .|.||.|||||.:.|.::.| || ||:||:: ||:.
Db 1478 PQDAVSRSQRRGRTGRGRMGIYRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAETSV 1537

Qy 1449 ILDTYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLQLCHQY 1507
        .| :.||||||| : :.|::||.:|  :.|.: :  :.||:.:||: ||.|||:. |:
Db 1538 RLRAYLNTPGLPVCQDHLEFWESVFTGLTHIDAHFLSQTRQAGDNFPYLVAYQATVCARA 1597

Qy 1508 GYAAPNDAPRWQG-ARLGKK------PCGVLWRLDGADACPGPEPSEVTRYQM---EVN 1559
        .   ::| :|.:  |.: :       |.|:..|| |..| || ::.|:  :    ||
Db 1598 QAPPPSWDQMWECLIRLKPTLHGPTPLLYRL-GAVQNEVTTTHPITKYIMACMSADLEVV 1656

Qy 1560 TSGTAALAVGVGVAMAYLAIDFFGATCVRRCWSITSVPTGATVAPVVDEEEI------VE 1613
        || |.:|.||:. |::||  . .. :.: :.:|:.||.||:.: |:. .        ::
Db 1657 TS-TWVLVGGVLAALAAYCLTTGSVVIVGR-------IILSGKPAIIPDREVLYREFDEME 1709

Qy 1614 ECASFIP-LEAMVAAIDKLK---------STIT----TTSPFTLETALBKLNTFLGPHAATI 1661
        ||||.:| :| |: |:::          |::     .  .:.::    .||:|| .|...:
Db 1710 ECASHLPYIEQGMQLAEQFKQKAIGLLQTATKQABAAAP--VVESKWRTLEAFWAKHMWNF 1768
```

```
Db  2782  :|||||||:  ||.|||||  ||:|||||  ||||||||  :|:|:||  ||||  |||    |:||:|||||  ||.||  ||:||  ||:|:|||  ||.||  :|.|:||
          ITSCSSNVSVAHDASGKRVYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLWAR  2841
Qy  2694  RVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEARFSVVRYTNAEI  2753
Db  2842  MILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGLSAFSLHSYSPGEI   2901
Qy  2754  LRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WHATSR----PLP-    2806
Db  2902  NRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWAVRTKLKLTPIPAA   2961
Qy  2807  ---DLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAI   2863
Db  2962  SQLDLSSWFVAGYS----------GDIYHSLSRARPRW-----FMWCLLLLSVGVGI     3004
```

```
Query Match       21.1%;  Score 3247;  DB 3;  Length 3010;
Best Local Similarity  30.2%;  Pred. No. 4.6e-270;
Matches  976;  Conservative 434;  Mismatches 1213;  Indels 612;  Gaps 114;

Qy     8  TSPVPAPRTRKNKQTQASYPVSIK----------------TSVERGQRAKRKVQRDARPR    51
Db     3  TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKASERSQPR   59
Qy    52  NYKIAGIHDGLQTLAQQALP--------------------------AHGNGRQD        79
Db    60  GRR--------QPIPKARRPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPTD   111
Qy    80  PRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG----  136
Db   112  PRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLP   170
Qy   137  WFGVHLFVVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC   195
Db   171  GCSFSIFLLALLSCLTIPASAYEVRNV-SGIYHVTNDCSNSSIVYEAADVIMHTPGCVPC   229
Qy   196  ADE------CWPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVG   250
Db   230  VQEGNSSRCWALTPTLA-ARNASVPTTTIRRHVDLLVGTRAAFCSAMYVGDLCGSIFLVS   288
Qy   251  DWLV----RHWLIHIDLNETGTCYLEVPTGIDPGFLG--FIGW----MAGKVEAVIPLTKL   301
Db   289  QLFTFSPRRHETVQ-DCN---------CSIYPGHVSGHRMAWDMMNMWSPTTALVVSQGI   337
Qy   302  ASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYI---------EAT       347
Db   338  L-RIPQAVVDMVAGAHWGVLAGLAYYSMVGNWAKVLIVALLFAGVDGETHTTGRVAGHTT   396
Qy   348  SGNPIRVPTGCS----------------------IAEFCSPLMIPCPCHSYLSENVS      382
```

```
Db   397  |||   ||||||||  :|  ||:|||  ||||||:.||  ||||:|||:  ||  |||:||.:|  ||:  |:|||||||||
          SGFTSLFSSGASQKIQLVNTNGSWHINRTALNCNDSLQTGFAALFY--AAHKFNSSGCP    453
Qy   383  E--VICYSPKWTR----PITLEYNNSISWYPY------------------------TIPGARGC-    416
Db   454  ERMASCRPIDWFAQGWGPITYTKPNSSDQRPYCWHYAPRPCGVVPASQVCGPVYCFTPSP   513
Qy   417  ------------------MVKFKNT----------WGCCRIRNVPSY----CTMG-      440
Db   514  VVVGTTDRSGVPTYSWGENETDVMLLNNTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGG   573
Qy   441  --------TDAVWNDTRNTYEACGVTPWLT-----------TAWHNGSALKLAILQ---YP   479
Db   574  VGNRTLICPTDCFRKHPEATYTKCGSGPWLTPRCLVDYPYRLWHYPCTLNFSIFKVRMYV   633
Qy   480  GSKE--MFKPHNWMSG-HLYPEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWL   536
Db   634  GGVEHRLNAACNWTRGRCNLEDRDR----------SELSPLLLSTTEWQILPCA------   678
Qy   537  QVPQGFVSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILGLCGSKYL   596
Db   679  --------FTTLPALSTGLIHLHQNIVDVQYLYGVGSAFSFAIKWEYILLFLLLADARVC   732
Qy   597  ILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLR   656
Db   733  ACLWMML:---------------------IAQAEAALENLVV--------------LN    755
Qy   657  YAALLGFVPMAAGLPLTFFVAAAAAQPDYDWMVRLLVAGLVLWAGRNRGHRIALLVGPWP   716
Db   756  AASVAG----AHGI-LSFLVFPCAA----WYIK--------GRLAPGAAYAFYGVWP    795
Qy   717  LVALLTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFPAHLLPRCALVNSYLWQRW   776
Db   796  L--LLLLLALPPRAYALDREMAASCGGAVLVGLVFLTLSPYYKVFLTRLIWWLQYFITRA   853
Qy   777  EN-----WFWNVTLRPERFFLVLV----CFPGATYDALVTFCVCHVALLCLTSSAAS---FF   826
Db   854  EAHMQVVMVPPLNVRGGRDAIILLTCAVHPELIFDITKLLLAILGPLMVLQAGITRVPYFV   913
Qy   827  GTDSRVRAHRMLVRLGKCHAWYSHYVLKFFLLVFGENGVFYKHL----------HGDVLPND   879
Db   914  RAQGLIRA-CMLVR------KVAGGHYVQMVFMKLGALTGTYVVNHLTPLRDWAHAGL--RD   966
Qy   880  FASKLPLQEP--FFPEFGKARVYRNEGRRLACGDTVDGLPVVARLGDLVFAGLA--MPPD   935
Db   967  LAVAV---EPVVFSAMETKVITWGAD--TAACGDIILGLPVSARRGKEIFLGPADSLEGQ   1021
Qy   936  GWAITAPFTLQCLSERGTLSAMAVVMVTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTA   995
Db  1022  GWRLLAPITAYSQQTRGVLGCIITSLTGRDKNQVEGEVQVVSTATQSPLATCINGVCWTV   1081
Qy   996  HHGSKGRRLAHPTGSIHPITVDAAN-DODI--YQPPCCAGSLTRCSSCGETKGYLVTRLGS   1052
Db  1082  YHGAGSKTLAGPKG---PITQMYTNVDLDLVGWQAPPGARSWTPCSCGSSSDLYLVTRHAD   1138
Qy  1053  LVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAA------RNSGGSVSQIRV   1109
Db  1139  VIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHVVGVFRAAVCTRGVAKAVDFIPV   1198
Qy  1110  RPLVCAGYHPQYTAHATLDTKPTVPNEYSVQLLIAPTGSGKSTKLPLSYMQEKYEVLVLN   1169
Db  1199  ESMETTMRSPVFTDNS---TPPAVPQTFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLN   1255
Qy  1170  PSVATTASMPKYMHATYGVNPNCYFNGKCCTNTGASLTYSTYGMYLT-GACSRN-YDVIIC   1227
Db  1256  PSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGGSIYSTYGKFLADGGCSGGAYDIIIC   1315
Qy  1228  DECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDEGTIPF   1287
Db  1316  DECHSTDSTTILGIGTVLDQAETAGARLVVLATATPGSVTVPHPNIEIEIGLSNNGEIPF   1375
Qy  1288  HGKKIKEENLKGKRHLIFEATKKHCDELANELARKGITAVSYYRGCDISKIPE-GDCVVV   1346
```

```
Qy   2286 FKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQVRLAK 2345
          :: .|||:.|| | |||  |  |::.:   :  : :.:.
Db   2432 PCAAEESKLPINPLSNSLLRHHNMVYATTSRSASLRQKKVTFDRLQVLDDHYRDVLKEMK 2491

Qy  °2346 EKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDL----QKCVEAG 2401
           ||:   .  :|| |
Db   2492 AKASTVKAKLLSIEEACKLTPPHSAKSKF-GYGAKDVRNLSSR-AVNHIRSVWEDLLEDT 2549

Qy   2402 EIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAVM 2461
          :  |:|:|||  .||| ||
Db   2550 ETPI---DTTIMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYDVVSTLPQAVM 2606

Qy   2462 GDAYGF-VDPRTRVKRLLSMWSPD--AVGATCDTVCFDSTITPEDIMVETDIYSAAKLSD 2518
          |.| |  |.  .:|:  |.|:
Db   2607 GSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTESDIRVBESIYQCCDLAP 2666

Qy   2519 QHRAGIHTIARQLYAGGPMIAYDGREIGYRCRCRSSGVVTTSSSNSLTCWLKVNAAAEQAG 2578
          ||:  |:  |||| |||| ||| |:||||  |::.|  ||
Db   2667 EARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKATAACRAAK 2726

Qy   2579 MKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCS 2638
          :|:  | |||:|| || |:: | :: ::.| |:: |:||||||
Db   2727 LQDCTMLVNGDDLVVICESAGTQEDAAALRAFTEAMTRYSAPGDPPQEYDLELITSCS 2786

Qy   2639 SNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAV 2698
          ||.|  .| :   | .|:|||| :|||.|| ::|| :  |||  |:  |
Db   2787 SNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPINSWLGNIIMYAPTLWARMILMT 2846

Qy   2699 HFMEQMLFEDKLPETVTFDWYGKNYTVPEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQ 2758
          || :. |.: || || |||  :: |||| |:|| ||| |  :  |||| |
Db   2847 HFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFTLHSYSPGEINRVAS 2906

Qy   2759 SLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WHATSR----PLP----DL 2808
          |:  .  ||.| |  | :  .:| :  |  |:
Db   2907 CLRKLGVPPLRTWRHRARSVRAKLLSQGGRAATCGRYLFNWAVRTKLKLTPIPAASQLDL 2966

Qy   2809 DKTSVARYTTFNYCDVVSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAI 2863
          |  |: ||
Db   2967 SGWFVAGYS----------GGDIYHSLSRARPRW------FPLCLLLLSVGVGI 3004
```

Search completed: October 27, 2005, 15:47:52
Job time : 93 secs

```
Db   1376 YGKAIPIBAIKGGRHLIFPCHSKKKKCDELAAKLTGLGLNAVAYYRGLDVSVIPPIGDVVVV 1435
Qy   1347 ATDALCTGYTGDFDSVYDCSLMVEGTCHVDLLDPTFTMGVRVCCGVSAIVKGQRRGRTGRGR 1406
          |||: |||||| ||: |  | | ||| | |  :
Qy   1436 ATDALMTGFTGDFDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGRTGRGR 1495

Qy   1407 AGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANL 1466
          :||  |||   |||| |:
Db   1496 SGIYRFVTPGERPSGMPDSSVLCECYDAGCAWYELTPAETSVRLRAYLNTFGLPVCQDHL 1555

Db   1467 DEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLQLCHQYGYAAPNDAPRWQG-ARLG 1524
          :|:   | |||| |||  |:: | :: ::.|
Qy   1556 EFWESVFTGLTHIDAHFLSQTKQAGDNFPYLVAYQATVCARAQAPPSWDQNMKCLIRLK 1615

Db   1525 KK---PCGVLWRLDGADACPGPEPSEV------TRYQM-CFT---EVNTSGTAALAVGVG 1571
          ||: |: |||| |||| |||
Qy   1616 PTLHGPTPLLYRL------GAVQNEVILTHPITKYIMACMSADLEVVTS-TWVLVGGVL 1667

Db   1572 VAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI--------VEECASFIP-LEAM 1624
          |::  |: |||| ||| |||  |:|||| |::.|  ||
Qy   1668 AALAAYCLTTGGSVVIVGR------IILLSGKPAVVPDREVLYQEFDEMEECASQLPYIEQG 1721

Db   1625 VAAIDKLK--------STIT----TTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVT 1673
          :|:  | |||:|| || |:: | :: ::.| |::
Qy   1722 MQLAEQFQKALGLLQTATKQAEAAAP-VVESKWRALETFWAKHMWNFISGIQYLAGLST 1780

Db   1674 LPDNPFASCVFAPIAGITTPLPHKIKMPLSLFGGAIASKLTDARGALAFMMAGAAGTALG 1733
          ||.|  .| :   | .|:|||| :|||.|| ::|| :  |||  |:  |
Qy   1781 LPGNPAIASLMAFTASITSPLTTQNTLLFNILGGWVAAQLAPPSAASAFVGAGIAGAAVG 1840

Db   1734 TWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVGVL 1790
          || :. |.: || || |||  :: |||| |:|| ||| |  :
Qy   1841 ---SIGLGKVLVDILAGYGAGVAGALVAFKVMSGEVFSTEDLVNLLPAILSPCALVVGVV 1897

Db   1791 SACAMFALTTAGPD----HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILLEASTPWS 1846
          |:  .  ||.| |  | :  .:| :  |
Qy   1898 --CAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAARVTQILSSLT--- 1952

Db   1847 VISACIRWLHTPTEDDCGLI---AWGLEIWQYVCNFFVICFNVLKAGVQS--MVNIPGCP 1901
          |  |: ||
Qy   1953 -ITQLLKRLHQWINEDCSTPCSGSWLRDVWDWICT----VLTDFKTWLQSKLLPRLPGVP 2007

Db   1902 FYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARL 1961
          |:||:|||| ||| |.|:  | |||||  ||  | : .:
Qy   2008 FLSCQRGYKGVWRGDGIMQTTCPCGAQIAGHVKNGSMRIV-GPRTCSNTWHGTFPINAYT 2066

Db   1962 CGSARPDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFTQVP--------- 2010
          .|.  |   ||||  |:|    |
Qy   2067 TGPCTPSPAPNYSRALMRVAAEEYVEVTRVGDFHYVTGMTTDNVKCPCQVPAPEFFTEVD 2126

Db   2011 ---------------PTLRAAVAVD------GVQVQCYLGEPKTPWTTSACCYGPDGKGKT 2050
          .|:  ||   | |: |||  .||
Qy   2127 GVRLHRYAPACKPLLREDVTFQVGLNQYLVGSSQLPC---EPE----------PDVTVLT 2172

Db   2051 VKL--PPRVDGHTPGVRMQLNLRDALETND----------CNSTNNTPSDEAAVSALV 2096
          .|:   | | ||: |  |   |:
Qy   2173 SMLTDPSHITAETAKRRLARGSPPSLASSSASQLSAPSLKATCTTHHDSPDADLIEANLL 2232

Db   2097 FKQEL------RRTNQLL-----EAISAGVDTTKLPAPSIEVVVRKRQFRARTGSLTL 2144
          ::||: |  |: |::
Qy   2233 WRQEMGGNITRVESENKVVILDSFEPLHAEGDEREISVAA--EILRKSRKFPSALPTWAR 2290

Db   2145 P---PP-----PRSVPGV--SCPESLQRSDPLEGPSNLPPSPP------VLQLAMPM 2185
          |   ||     |::   |:   |
Qy   2291 PDYNPPLLESWKDPDYVPPVVHGCPLP--------PTKAPPIPPPRRKRTVVLTESNVS 2341

Db   2186 PLLGAGECNPFTAIGCAMTETGGG---PD-----DLPSY---PP---KKEVSEWS 2226
          .|   |:  .| |::|    ||    |   |
Qy   2342 SALAELATKTFGSSGSSAVDSGTATALPDLASDDGDKGSDVESYSSMPPLEGEPGDPDLS 2401

Db   2227 DESWSTATTASSVTGPPYPYPKIRGKDSTQSAPAKRPFTKKKLGKSEFSCSMSYTWTD-VIS 2285
          |.|||| ||  |            |::  |     ||||||||:  ||:
Qy   2402 DGSWSTVSEEAS------------EDVVCCSMSYTWTGALIT 2431
```

GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2005, 15:36:22 ; Search time 286 Seconds
                                    (without alignments)
                        4186.277 Million cell updates/sec

Title: US-10-009-002-2
Perfect score: 15376
Sequence: 1 MPVISTQTSPVPAPRTRKNK.........KYLAVIVFALGLIAVGLAIS 2864

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1865214 seqs, 41804304O residues

Total number of hits satisfying chosen parameters: 1865214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published Applications AA:*
         1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
         2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
         3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
         4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
         5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
         6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
         7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
         8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
         9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
        10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
        11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
        12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
        13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
        14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
        15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
        16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
        17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
        18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
        19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
        20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
        21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
        22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	15376	100.0	2864	9	US-09-742-659-2	Sequence 2, Appli
2	15317	99.6	2864	8	US-08-424-550B-394	Sequence 394, App
3	14474	94.1	2862	9	US-09-742-659-5	Sequence 5, Appli
4	13194	85.8	2865	9	US-09-742-659-6	Sequence 6, Appli
5	7380	48.0	1422	8	US-08-424-550B-83	Sequence 83, Appli
6	3262	21.2	3010	15	US-10-467-000-1	Sequence 1, Appli
7	3234.5	21.0	3011	15	US-10-296-734-406	Sequence 406, App
8	3232	21.0	3010	16	US-10-333-449A-34	Sequence 34, Appl
9	3219.5	20.9	3011	9	US-09-742-659-4	Sequence 4, Appli
10	3215.5	20.9	3011	20	US-11-126-662-1	Sequence 1, Appli
11	3214.5	20.9	3011	9	US-09-952-572-9	Sequence 9, Appli
12	3214.5	20.9	3011	9	US-09-747-419-20	Sequence 20, Appl
13	3214.5	20.9	3011	14	US-10-259-275-20	Sequence 20, Appl
14	3214.5	20.9	3011	15	US-10-189-359-14	Sequence 14, Appl
15	3214.5	20.9	3011	20	US-11-006-313-20	Sequence 20, Appl
16	3208.5	20.9	3011	9	US-09-891-894-3	Sequence 3, Appli
17	3208.5	20.9	3011	14	US-10-184-150-3	Sequence 3, Appli
18	3208.5	20.9	3011	15	US-10-328-997-3	Sequence 3, Appli
19	3208.5	20.9	3012	9	US-09-238-076-2	Sequence 2, Appli
20	3208.5	20.9	3012	10	US-09-995-937-2	Sequence 2, Appli
21	3208.5	20.9	3012	10	US-09-917-563-2	Sequence 2, Appli
22	3204.5	20.9	3011	9	US-09-916-359-2	Sequence 2, Appli
23	3204.5	20.8	3011	16	US-10-445-724-2	Sequence 2, Appli
24	3202.5	20.8	3011	9	US-09-238-076-20	Sequence 20, Appl
25	3202.5	20.8	3011	10	US-09-995-937-20	Sequence 20, Appl
26	3202.5	20.8	3011	10	US-09-917-563-20	Sequence 20, Appl
27	3188	20.7	2940	14	US-10-226-629A-13	Sequence 13, Appl
28	3148.5	20.5	2985	14	US-10-259-275-40	Sequence 40, Appl
29	3148.5	20.5	2985	20	US-11-006-313-40	Sequence 40, Appl
30	3129.5	20.4	3011	14	US-10-232-643-6	Sequence 6, Appli
31	3126.5	20.3	3011	9	US-09-929-955-1	Sequence 1, Appli
32	3126.5	20.3	3011	13	US-10-104-966-1	Sequence 1, Appli
33	3126.5	20.3	3011	15	US-10-719-619-1	Sequence 1, Appli
34	3126.5	20.3	3011	16	US-10-817-591-1	Sequence 1, Appli
35	3045	19.8	2894	9	US-09-941-611-23	Sequence 23, Appl
36	3045	19.8	2894	14	US-10-044-995-23	Sequence 23, Appl
37	3045	19.8	2894	17	US-10-822-871-23	Sequence 23, Appl
38	2992	19.5	576	13	US-10-211-455-3	Sequence 3, Appli
39	2830.5	18.4	2201	13	US-10-029-907-3	Sequence 3, Appli
40	2830.5	18.4	2201	14	US-10-309-561-3	Sequence 3, Appli
41	2830.5	18.4	2201	16	US-10-789-355-3	Sequence 3, Appli
42	2830.5	18.4	2201	16	US-10-686-835-3	Sequence 3, Appli
43	2795.5	18.2	2201	13	US-10-085-476-2	Sequence 2, Appli
44	2754.5	17.9	1985	14	US-10-259-275-42	Sequence 42, Appl
45	2754.5	17.9	1985	16	US-10-639-150-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-742-659-2

; Sequence 2, Application US/09742659
; Patent No. US200100034019A1
; GENERAL INFORMATION:
; APPLICANT: Hong, Zhi
; APPLICANT: Butkiewicz, Nancy J.
; APPLICANT: Zhong, Weidong
; APPLICANT: Ingravallo, Paul
; APPLICANT: Wright-Minogue, Jacquelyn
; APPLICANT: Lau, Johnson Y.
; APPLICANT: Lemon, Stanley M.
; TITLE OF INVENTION: Chimeric HCV/GBV-B viruses
; FILE REFERENCE: ID01116
; CURRENT APPLICATION NUMBER: US/09/742,659
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/171,469
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 2
; LENGTH: 2864
; TYPE: PRT
; ORGANISM: GB virus-B
US-09-742-659-2

Query Match   100.0%; Score 15376; DB 9; Length 2864;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 2864; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD 60

```
Qy    61  GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTTHTPLVGAVVRPV  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTTHTPLVGAVVRPV  120

Qy   121  CQIVRLLEDGVNATGWFGVHLFVVCLLSLACPCSGARVTDPDNTTLILTNCCQRNQVIY  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  CQIVRLLEDGVNATGWFGVHLFVVCLLSLACPCSGARVTDPDNTTLILTNCCQRNQVIY  180

Qy   181  CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG  240

Qy   241  ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK  300

Qy   301  LASQVPYAIATMFSSVHYLAVGALLYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  LASQVPYAIATMFSSVHYLAVGALLYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI  360

Qy   361  AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF  420

Qy   421  KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG  480

Qy   481  SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ  540

Qy   541  GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVLILLGLCGSKYLILAY  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVLILLGLCGSKYLILAY  600

Qy   601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL  660

Qy   661  LGFVPMAAGLPLTFFVAAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALLVGPWPLVAL  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  LGFVPMAAGLPLTFFVAAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALLVGPWPLVAL  720

Qy   721  LTLLHLVTPASAFDTEIIGGLTIPPVALVVMSRFGFFAHLLPRCALVNSYLWQRWENWF  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  LTLLHLVTPASAFDTEIIGGLTIPPVALVVMSRFGFFAHLLPRCALVNSYLWQRWENWF  780

Qy   781  WNVTLRPERFFLVLVCFPGATYDALVTFCVCHVALLCLTSSAASFGTDSRVRAHRMLVR  840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  WNVTLRPERFFLVLVCFPGATYDALVTFCVCHVALLCLTSSAASFGTDSRVRAHRMLVR  840

Qy   841  LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPFEGKARVY  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPFEGKARVY  900

Qy   901  RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPFTLQCLSERGTLSAMAVV  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPFTLQCLSERGTLSAMAVV  960

Qy   961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020

Qy  1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG 1080

Qy  1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQTAHATLDTKPTVPNEYSVQ 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQTAHATLDTKPTVPNEYSVQ 1140

Qy  1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200
```

```
Db  1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1201  TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT 1260

Qy  1261  ATPPGVIPTPHANTEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKGHCDELANELA 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  ATPPGVIPTPHANTEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKGHCDELANELA 1320

Qy  1321  RKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  RKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT 1380

Qy  1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYG 1440

Qy  1441  LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ 1500

Qy  1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT 1560

Qy  1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEBIVEECASFIP 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEBIVEECASFIP 1620

Qy  1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA 1680

Qy  1681  SCVFAPIAGITTLPLPHKIKMFLSFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  SCVFAPIAGITTLPLPHKIKMFLSFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF 1740

Qy  1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT 1800

Qy  1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE 1860

Qy  1861  DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  DDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ 1920

Qy  1921  ARCPCCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVNYG 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  ARCPCCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVNYG 1980

Qy  1981  VRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVDGVQVCYLGEPKTPWTTSAC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  VRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVDGVQVCYLGEPKTPWTTSAC 2040

Qy  2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE 2100

Qy  2101  LRRTNQLLEAISAGVDTTKLPAPSIEVVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL 2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  LRRTNQLLEAISAGVDTTKLPAPSIEVVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL 2160

Qy  2161  QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK 2220

Qy  2221  EVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW 2280
```

```
Db   EVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280
Qy   TDVISFKTASKVLSATRAITSGFLKQRSLVYYVTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340
Db   TDVISFKTASKVLSATRAITSGFLKQRSLVYYVTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340
Qy   VRLAKEKASKVGVGMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400
Db   VRLAKEKASKVGVGMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400
Qy   GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV  2460
Db   GEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV  2460
Qy   MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520
Db   MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520
Qy   RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK  2580
Db   RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEQAGMK  2580
Qy   NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN  2640
Db   NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN  2640
Qy   VTSGITKSGKPYYFLTRDPRIPLGRCSAEBGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF  2700
Db   VTSGITKSGKPYYFLTRDPRIPLGRCSAEBGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF  2700
Qy   MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL  2760
Db   MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL  2760
Qy   TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN  2820
Db   TDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN  2820
Qy   YCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS  2864
Db   YCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS  2864
```

```
CURRENT APPLICATION DATA:
   APPLICATION NUMBER:  US/08/424,550B
   FILING DATE:
   CLASSIFICATION:  435435
ATTORNEY/AGENT INFORMATION:
   NAME:  POREMBSKI, PRISCILLA E.
   REGISTRATION NUMBER:  33,207
   REFERENCE/DOCKET NUMBER:  5527.PC.01
TELECOMMUNICATION INFORMATION:
   TELEPHONE:  708-937-6365
   TELEFAX:  708-938-2623
INFORMATION FOR SEQ ID NO:  394:
SEQUENCE CHARACTERISTICS:
   LENGTH:  2864 amino acids
   TYPE:  amino acid
   TOPOLOGY:  linear
   MOLECULE TYPE:  protein
US-08-424-550B-394

Query Match           99.6%;  Score 15317;  DB 8;  Length 2864;
Best Local Similarity 99.6%;  Pred. No. 0;
Matches 2852;  Conservative 4;  Mismatches 8;  Indels 0;  Gaps 0;
```

```
Qy    1  MPVISTQTSPVPAPRTRKNKQTQASYPVS1KTSVERGQRAKRKVQRDARPRNYKIAGIHD  60
Db    1  MPVISTQTSPVPAPRTRKNKQTQASYPVS1KTSVERGQRAKRKVQRDARPRNYKIAGIHD  60
Qy   61  GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPV 120
Db   61  GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPV 120
Qy  121  CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180
Db  121  CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180
Qy  181  CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240
Db  181  CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240
Qy  241  ELCGACVLVGDWLVRHWLHIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK 300
Db  241  ELCGACVLVGDWLVRHWLHIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK 300
Qy  301  LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI 360
Db  301  LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI 360
Qy  361  AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF 420
Db  361  AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPVTLEYNNSISWYPYTIPGARGCMVKF 420
Qy  421  KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAIIQYPG 480
Db  421  KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAIIQYPG 480
Qy  481  SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ 540
Db  481  SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ 540
Qy  541  GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLCGSKYLILAY 600
Db  541  GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLCGSKYLILAY 600
Qy  601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL 660
Db  601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL 660
Qy  661  LGFVPMAAGLPLTFFVAAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALLVGPMLVAL 720
Db  661  LGFVPMAAGLPLTFFVAAAAAQPDYDWWVRLLVAGLVLWAGRDRGPRIALLVGPMLVAL 720
Qy  721  LTLHLVTPASAFDTEIIGGLTIPPVALVVMSRGFFAHLLPRCALVNSYLWQRWENWF 780
```

```
RESULT 2
US-08-424-550B-394
; Sequence 394, Application US/08424550B
Publication No. US20020119447A1
GENERAL INFORMATION:
   APPLICANT:  JOHN N. SIMONS
   APPLICANT:  TAMI J. PILOT-MATIAS
   APPLICANT:  GEORGE J. DAWSON
   APPLICANT:  GEORGE G. SCHLAUDER
   APPLICANT:  SURESH M. DESAI
   APPLICANT:  THOMAS P. LEARY
   APPLICANT:  ANTHONY SCOTT MUERHOFF
   APPLICANT:  JAMES C. ERKER
   APPLICANT:  SHERI L. BUIJK
   APPLICANT:  ISA K. MUSHAHWAR
   TITLE OF INVENTION:  NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
   TITLE OF INVENTION:  REAGENTS AND METHODS FOR THEIR USE
   NUMBER OF SEQUENCES:  716
   CORRESPONDENCE ADDRESS:
      ADDRESSEE:  ABBOTT LABORATORIES D377/AP6D
      STREET:  100 ABBOTT PARK ROAD
      CITY:  ABBOTT PARK
      STATE:  IL
      COUNTRY:  USA
      ZIP:  60064-3500
   COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
Db    721  LTLLHLATPASAFDTEIIGGLTIPPVALVVMSRFGFFAHLLPRCALVNSYLWQRWENWF   780

Qy    781  WNVTLRPERRFFLVLVCFPGATYDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR  840
           |||||||||| |||||||||| |||||||||||||||||||||||||||||||||||||||
Db    781  WNVTLRPERRFLLLVCFPGATYDTLVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR   840

Qy    841  LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPFEGKARVY   900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPFEGKARVY   900

Qy    901  RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPFTLQCLSERGTLSAMAVV   960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPFTLQCLSERGTLSAMAVV   960

Qy    961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN  1020

Qy   1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG  1080

Qy   1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ  1140

Qy   1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN  1200

Qy   1201  TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT  1260

Qy   1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIPEATKHCDELANELA   1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIPEATKKHCDELANELA  1320

Qy   1321  RKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  RKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPT  1380

Qy   1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVVDGSCTPSGMVPECNIVEAFDAAKAWYG  1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVVDGSCTPSGMVPECNIVEAFDAAKAWYG  1440

Qy   1441  LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ  1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441  LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ  1500

Qy   1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT  1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT  1560

Qy   1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDBEEIVEECASFIP  1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDBEEIVEECASFIP  1620

Qy   1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATLAIIEYCCGLVTLPDNPFA   1680
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATLAIIEYCCGLVTLPDNPFA  1680

Qy   1681  SCVFDAGIAGTTTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF  1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681  SCVFAGIAGTTTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF  1740

Qy   1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVGVLSACAMFALTT   1800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVGVLSACAMFALTT  1800

Qy   1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKIIGILEASTPWSVISACIRWLHTPTE  1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKIIGILEASTPWSVISACIRWLHTPTE  1860
```

```
Qy   1861  DDCGLIAWGLEIWQYVCNFVICFNVLKAGVQSMNIPGGPFYSCQKGYKGPWIGSGMLQ   1920
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861  DDCGLIAWGLEIWQYVCNFVICFNVLKAGVQSMNIPGGPFYSCQKGYKGPWIGSGMLQ   1920

Qy   1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVNYG  1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVNYG  1980

Qy   1981  VRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVADGVQVQCYLGEPKTPWTTSAC   2040
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981  VRDYCKYEKLGDHIFVTAVSSPNVCFTQVPPTLRAAVADGVQVQCYLGEPKTPWTTSAC  2040

Qy   2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE  2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSINNTPSDEAAVSALVFKQE  2100

Qy   2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL   2160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPPRSVPGVSCPESL  2160

Qy   2161  QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK  2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161  QRSDPLEGPSNLPSSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPPKK  2220

Qy   2221  EVSEWSDESWSTATTASSYYTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221  EVSEWSDGSWSTTTTASSYYTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW  2280

Qy   2281  TDVISFKTASKVLSATRAITSGFLKQRSLVYTEPRDAELRKQKVTINRQPLFPPSYHKQ   2340
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281  TDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQ  2340

Qy   2341  VRLAKEKASKVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA   2400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341  VRLAKEKASKVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEA  2400

Qy   2401  GEIPSHYRQTVIVPKEEVFVKTPQKPKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV   2460
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401  GEIPSHYRQTVIVPKEEVFVKTPQKPKKPPRLISYPHLEMRCVEKMYYGQVAPDVVKAV  2460

Qy   2461  MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461  MGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQH  2520

Qy   2521  RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEEQAGMK  2580
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521  RAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAAEEQAGMK  2580

Qy   2581  NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN  2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581  NPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSN  2640

Qy   2641  VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF  2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641  VTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHF  2700

Qy   2701  MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL  2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701  MEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSL  2760

Qy   2761  TDMTMPPLRAWRKKARAVLASAKRRCGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN  2820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761  TDMTMPPLRAWRKKARAVLASAKRRCGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFN  2820

Qy   2821  YCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS  2864
           ||||| |||||||||||||||||||||||||||||||||||||||
Db   2821  YCDVYSPEGDVFVTPQRRLQKFLVKYLAVIVFALGLIAVGLAIS  2864
```

RESULT 3
US-09-742-659-5
; Sequence 5, Application US/09742659

Patent No. US20010034019A1
GENERAL INFORMATION:
; APPLICANT:
; APPLICANT: Hong, Zhi
; APPLICANT: Butkiewicz, Nancy J.
; APPLICANT: Zhong, Weidong
; APPLICANT: Ingravallo, Paul
; APPLICANT: Wright-Minogue, Jacquelyn
; APPLICANT: Lau, Johnson Y.
; APPLICANT: Lemon, Stanley M.
; TITLE OF INVENTION: Chimeric HCV/GBV-B viruses
; FILE REFERENCE: ID01116
; CURRENT APPLICATION NUMBER: US/09/742,659
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/171,469
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 5
; LENGTH: 2862
; TYPE: PRT
; ORGANISM: GBV-B/HCV
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (945)..(1129)
; OTHER INFORMATION: chimeric region
; NAME/KEY: SITE
; LOCATION: (1579)..(1593)
; OTHER INFORMATION: chimeric region
US-09-742-659-5

Query Match   94.1%;   Score 14474;   DB 9;   Length 2862;
Best Local Similarity 94.7%;   Pred. No. 0;
Matches 2717;   Conservative 39;   Mismatches 101;   Indels 12;   Gaps 5;

Qy   1 MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD 60

Qy   61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPV 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPV 120

Qy   121 CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180

Qy   181 CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240

Qy   241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTK 300

Qy   301 LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYIEATSGNPIRVPTGCSI 360

Qy   361 AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF 420

Qy   421 KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG 480

Qy   481 SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQ 540

Qy   541 GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILAY 600

Qy   601 LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYLRCRLRYAAL 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFFICCYLRCRLRYAAL 660

Qy   661 LGFVPMAAGLPLTFVAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALLVGPWPLVAL 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 LGFVPMAAGLPLTFVAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALLVGPWPLVAL 720

Qy   721 LTLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQRWENWF 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 LTLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQRWENWF 780

Qy   781 WNVTLRPBRRFFLVLVCFPGATYDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 WNVTLRPBRRFFLVLVCFPGATYDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR 840

Qy   841 LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPFEGKARVY 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPFEGKARVY 900

Qy   901 RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPFTLQCLSERGTLSAMAVV 960
           ||||||||||||||||||||||||||||||||||||||: :
Db   901 RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPITAYAQQTRGLLGCIITS 960

Qy   961 MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020
           |:|| |:   :|| :.. :   : . . IQMYTNV 1018
Db   961 LTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRTIASPKGPV--IQMYTNV 1018

Qy   1021 DQDI--YQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGS 1078
           ||:| :::| ::||.   : :| :| ::| .   : :   :: :||
Db   1019 DQDLVGWPAPQGSRSLTPCTCGGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGS 1078

Qy   1079 SGAPILCSSGHVIGMFTAA---RNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPN 1135
           |.||: :|:| |::.:   :: ||.|::: |:||| : : ::SKPTVPN 1135
Db   1079 SGGPLLCPAGHAVGLFRAAVCTRGVAKAVDFIPVENLETMRSPVFTDNS---SKPTVPN 1135

Qy   1136 EYSVQILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFN 1195
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1136 EYSVQILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFN 1195

Qy   1196 GKCTNTGASLTYSYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRL 1255
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1196 GKCTNTGASLTYSYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRL 1255

Qy   1256 VVLATATPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDEL 1315
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1256 VVLATATPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDEL 1315

Qy   1316 ANELARKGITAVSYYRGCDISKIPEGDCVVVATDALCGYTGDFDSVYDCSLMVEGTCHV 1375
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1316 ANELARKGITAVSYYRGCDISKIPEGDCVVVATDALCGYTGDFDSVYDCSLMVEGTCHV 1375

Qy   1376 DLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYVVDGSCTPSGMVPECNIVEAFDAA 1435
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1376 DLDPTFTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYVVDGSCTPSGMVPECNIVEAFDAA 1435

Qy   1436 KAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVL 1495
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1436 KAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVL 1495

Qy   1496 LTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCF 1555
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1496 LTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCF 1555

Qy   1556 TEVNTSGTAALAVGVANAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEIVEEC 1615
           ||||||||||||||||||||||| :   : :|| ||||||||||||||
Db   1556 TEVNTSGTAALAVGVGVANAYLA--STGCVVIVGRIVLSGKPTGATVIVLSGKPTGATVAPVVDEEIVEEC 1613

Qy   1616 ASFIPLEAMVAAIDLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLP 1675
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1614 ASFIPLEAMVAAIDLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLP 1673

```
Db  2754 VSQSLTDMTMPPLRAWRKKARAVLASAKRGGAHAKLARFLLWHATSRPLPDLDKTSVAR 2813
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  2816 YTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS 2864
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  2814 YTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS 2862


RESULT 4
US-09-742-659-6
; Sequence 6, Application US/09742659
; Patent No. US20010034019A1
; GENERAL INFORMATION:
;   APPLICANT:  Hong, Zhi
;   APPLICANT:  Butkiewicz, Nancy J.
;   APPLICANT:  Zhong, Weidong
;   APPLICANT:  Ingravallo, Paul
;   APPLICANT:  Wright-Minogue, Jacquelyn
;   APPLICANT:  Lau, Johnson Y.
;   APPLICANT:  Lemon, Stanley M.
; TITLE OF INVENTION: Chimeric HCV/GBV-B viruses
; FILE REFERENCE: ID01116
; CURRENT APPLICATION NUMBER: US/09/742,659
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/171,469
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2865
; TYPE: PRT
; ORGANISM: GBV-B/HCV
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (2275)..(2865)
; OTHER INFORMATION: chimeric region
US-09-742-659-6

Query Match             85.8%;  Score 13194;  DB 9;  Length 2865;
Best Local Similarity   86.5%;  Pred. No. 0;
Matches  2496;  Conservative   98;  Mismatches  241;  Indels  52;  Gaps  12;

Qy    1 MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKVQRDARPRNYKIAGIHD 60

Qy   61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTHTPLVGPLVGAVVRPV 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVTHTPLVGPLVGAVVRPV 120

Qy  121 CQIVRLLEDGVNWATGWFGWLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CQIVRLLEDGVNWATGWFGWLFVVCLLSLACPCSGARVTDPDTNTTILTNCCQRNQVIY 180

Qy  181 CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIG 240

Qy  241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFGFIGWMAGKVEAVIFLTK 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTGIDPGFGFIGWMAGKVEAVIFLTK 300

Qy  301 LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLALMLYIEATSGNPIRVPTGCSI 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 LASQVPYAIATMFSSVHYLAVGALIYYASRGKWYQLLALMLYIEATSGNPIRVPTGCSI 360

Qy  361 AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AEFCSPLMIPCPCHSYLSENVSEVICYSPKWTRPITLEYNNSISWYPYTIPGARGCMVKF 420

Qy  421 KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG 480
```

```
Qy 1676 DNPPASCVPAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTW 1735
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1674 DNPPASCVPAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGTALGTW 1733

Qy 1736 TSVGFVPDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAM 1795
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1734 TSVGFVPDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAM 1793

Qy 1796 FALTTAGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWL 1855
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1794 FALTTAGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWL 1853

Qy 1856 HTPTEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIG 1915
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1854 HTPTEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIG 1913

Qy 1916 SGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSL 1975
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1914 SGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSL 1973

Qy 1976 VVNYGVRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDGVQVQCYLGEPKTPW 2035
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1974 VVNYGVRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDGVQVQCYLGEPKTPW 2033

Qy 2036 TTSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSAL 2095
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2034 TTSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSAL 2093

Qy 2096 VFKQELRRTNQLLEAISAGVDTTKLPAPSIEEVVRKQFRARTGSLTLPPPRSVPGVS 2155
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2094 VFKQELRRTNQLLEAISAGVDTTKLPAPSIEEVVRKQFRARTGSLTLPPPRSVPGVS 2153

Qy 2156 CPESLQRSDPLEGPSNLPFSPPVLQLAMEPMPLLGAGECNPFTAIGCAMTETGGGPDDLPS 2215
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2154 CPESLQRSDPLEGPSNLPFSPPVLQLAMEPMPLLGAGECNPFTAIGCAMTETGGGPDDLPS 2213

Qy 2216 YPPKKEVSEWSDESWSTATTASSVVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCS 2275
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2214 YPPKKEVSEWSDESWSTATTASSVVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCS 2273

Qy 2276 MSYTWTDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPP 2335
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2274 MSYTWTDVISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPP 2333

Qy 2336 SYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQ 2395
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2334 SYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQ 2393

Qy 2396 KCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPD 2455
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2394 KCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYYGQVAPD 2453

Qy 2456 VVKAVMGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAK 2515
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2454 VVKAVMGDAYGFVDPRTRVKRLLSMWSPDAVGATCDTVCFDSTITPEDIMVETDIYSAAK 2513

Qy 2516 LSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTSSSNSLTCWLKVNAAAE 2575
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2514 LSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTSSSNSLTCWLKVNAAAE 2573

Qy 2576 QAGMKNPRFLICGGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELT 2635
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2574 QAGMKNPRFLICGGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELT 2633

Qy 2636 SCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRV 2695
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2634 SCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRV 2693

Qy 2696 LAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILR 2755
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2694 LAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILR 2753

Qy 2756 VSQSLTDMTMPPLRAWRKKARAVLASAKRGGAHAKLARFLLWHATSRPLPDLDKTSVAR 2815
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2756 KNNTWGCCRIRNVPSYCTMGTDAVWNDTRNTYEACGVTPWLTTAWHNGSALKLAILQYPG 2815
```

```
Qy  481  SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPERWARLPGTPPVVRGSWLQVPQ  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  SKEMFKPHNWMSGHLYFEGSDTPIVYFYDPVNSTLLPERWARLPGTPPVVRGSWLQVPQ  540

Qy  541  GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILGLCGSKYLILAY  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  GFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILGLCGSKYLILAY  600

Qy  601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  LCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAAL  660

Qy  661  LGFVPMAAGLPLTFFVAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALLVGPWPLVAL  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  LGFVPMAAGLPLTFFVAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALLVGPWPLVAL  720

Qy  721  LTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLMQRWENWF  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  LTLLHLVTPASAFDTEIIGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLMQRWENWF  780

Qy  781  WNVTLRPERFFLVLVCFPGATVDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  WNVTLRPERFFLVLVCFPGATVDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVR  840

Qy  841  LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPPEGKARVY  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  LGKCHAWYSHYVLKFFLLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPPEGKARVY  900

Qy  901  RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPFTLQCLSERGTLSAMAVV  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  RNEGRRLACGDTVDGLPVVARLGDLVFAGLAMPPDGWAITAPFTLQCLSERGTLSAMAVV  960

Qy  961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  MTGIDPRTWTGTIFRLGSLATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAAN 1020

Qy 1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  DQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSG 1080

Qy 1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081  APILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQ 1140

Qy 1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141  ILIAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTN 1200

Qy 1201  TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201  TGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLAT 1260

Qy 1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261  ATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELA 1320

Qy 1321  RKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVVDCSLMVEGTCHVDLDPT 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321  RKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVVDCSLMVEGTCHVDLDPT 1380

Qy 1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYG 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381  FTMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYG 1440

Qy 1441  LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441  LSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYVLLTAAQ 1500

Qy 1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501  LQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQMCFTEVNT 1560
```

```
Qy 1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEBIVEECASFIP 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561  SGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEBIVEECASFIP 1620

Qy 1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621  LEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAIIEYCCGLVTLPDNPFA 1680

Qy 1681  SCVFAPIAGITTPLPHKIKMFLSLFPGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681  SCVFAPIAGITTPLPHKIKMFLSLFPGGAIASKLTDARGALAFMMAGAAGTALGTWTSVGF 1740

Qy 1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741  VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVGVLSACAMFALTT 1800

Qy 1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801  AGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTPWSVISACIRWLHTPTE 1860

Qy 1861  DDCGLIAWGLEIWQYVCNFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861  DDCGLIAWGLEIWQYVCNFVICFNVLKAGVQSMVNIPGCPFYSCQKGYKGPWIGSGMLQ 1920

Qy 1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVNYG 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921  ARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGSARPDPTDWTSLVVNYG 1980

Qy 1981  VRDYCKYERMGDHIFVTAVSSPNVCFTQVPPTLRAAVADVGVQOVCYLGEPKTPWTTSAC 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981  VRDYCKYERMGDHIFVTAVSSPNVCFTQVPPTLRAAVADVGVQOVCYLGEPKTPWTTSAC 2040

Qy 2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041  CYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNNTPSDEAAVSALVFKQE 2100

Qy 2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPRSVPGVSCPESL 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101  LRRTNQLLEAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLPPPRSVPGVSCPESL 2160

Qy 2161  QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPKKK 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161  QRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTETGGGPDDLPSYPKKK 2220

Qy 2221  EVSEWSDESWSTATTASSVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221  EVSEWSDESWSTATTASSVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTW 2280

Qy 2281  TD-VISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPSSYHK 2339
         ||  : :  ||||  ::  : :||   |:  |: ::   : ||| :||:|: :
Db 2281  TGALITPCAEBESKLPINALSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVLDDHYRD 2340

Qy 2340  QVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDQKCVE- 2399
         ||: ||||||| :  ||||  ||  ::   |||||: -|| |||  |    |||
Db 2341  VLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKF-GYGAKDVRN-LSSKAV-------- 2390

Qy 2400  AGEIPSHYR---------------QTVIVPKEEVVKTPQKPRRLISYPHLEMRCVEKMYY 2449
         ||: ||:||        : |||  ||||  ||| ||:  :  |||| :||:  |||:||
Db 2391  -NHIHSVWKDLLEDTVTPIDTTIMAKNEVFCVQPEKGGRKPARLLVFPDLGVRVCEKMAL 2449

Qy 2450  GQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMW--SPDAVGATCDTVCFDSTITPEDIMV 2506
         |:|  : |:  ||  :||  : ::||  ::|    ||   |  :| ||||||:| |:||:|
Db 2450  YDVVSTLPQVVMGSSYGFQYSPGQRVEFLVNTWKSKKNPMGFSYDTRCFDSTVTENDIRV 2509

Qy 2507  ETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDVAYDGREIGYRRCRSSGVYTTSSSNSLTC 2566
         |  ||   : |  ::||| |: |:|||||| :  | :  | |:| |||| ||| ||:  ||:||
Db 2510  EESIYQCCDLAPEARQAIKSLTERLYIGGPLTNSKGQNCCGYRRCRASGVLTTSCGNTLTC 2569

Qy 2567  WLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQ 2626
         :||: || |:    |  |:  ||||| ||| ||| :| :|::||:   :|  || | ||
Db 2570  YLKASAACRAAKLQDCTMLVNGDDLVVICESAGVQEDAASLRAFTEAMTRYSAPPGDPPQ 2629

Qy 2627  PKYSLEBLTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHH 2686
```

```
Db   2630 PEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIMY 2689
          |:|:||  |:|||||||:   |||    |:| ||||||     |  |  |:|:|
Qy   1005 AHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYW 1064

Db     60 ..........
Qy   1005 AHPTGSIHPITVDAANDQDIYQPPCGAGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYW 1064

Qy   2687 YPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVV 2746
          |:|   ::|| ||:|  |     :|::  :|  |::
Db     60 ........................................................... 60

Db   2690 APTLWARMLMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIERLHGLSAFSLH 2749
          || ||  :||   | :||    ||
Qy   1065 CVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAH 1124

Qy   1065 CVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVGQIRVRPLVCAGYHPQYTAH 1124
          |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db     61 CVCGALPMAVAKGSSGAPILCSSGHVIGMFTAARNSGGSVGQIRVRPLVCAGYHPQYTAH 120

Db   2747 RYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WHATSR-- 2803
          ::  |     |:
Qy   1125 ATLDTKPTVPNEYSVQILLAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHA 1184

Qy   1125 ATLDTKPTVPNEYSVQILLAPTGSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHA 1184
          ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db    121 ATLDTKPTVPNEYSVQILLAPTGSGKSTKLPLSYMQKKXEVLVLNPSVATTASMPKYMHA 180

Db   2750 SYSPGEINRVASCLRKLGVPLRVWRHRARSVRARLLSQGRAATCGKYLFNWAVKTKLK 2809
          :: |
Qy   1185 TYGVNPNCYFNGKCNTGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKV 1244

Qy   1185 TYGVNPNCYFNGKCNTGASLTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKV 1244
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TYGVNPNCYFNGKCNTGASLTYSTYGMYLTGRCSRNYDVIICDECHATDRTTVLGIGKV 240

Db   2804 --PLP-----DLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFLVXLAVIVFALGL 2856
          |:|:|
Qy   1245 LTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKENLKKGRHLI 1304

Qy   1245 LTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKENLKKGRHLI 1304
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 LTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDEGTIPFHGKKIKENLKKGRHLI 300

Db   2810 LTPIPAASQDLSGMFVAGYS----------GGDIYHSLSRARPRW------FMLCLLL 2852
          ::|:|
Qy   1305 FEATKKGHCDELANELARKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYD 1364

Qy   1305 FEATKKGHCDELANELARKGITAVSYYRGCDISKIPEGDCVVVATDALCTGYTGDFDSVYD 1364
          ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db    301 FEATKKGHCDELANELARKGITAVSYYRGCDISKMPEGDCVVVATDALCTGYTGDFDSVYD 360

Qy   2857 IAVGLAI 2863
          ::|:|:
Db   2853 LSVGVGI 2859

Qy   1365 CSLMVEGTCHVDLDPTFTMGRVCGVSAIVKGQRRGRTGRGAGIYYVVDGSCTPSGMVP 1424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CSLMVEGTCHVDLDPTFTMGRVCGVSAIVKGQRRGRTGRGAGIYYVVDGSCTPSGMVP 420

Qy   1425 ECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVN 1484
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVN 480

Qy   1485 TAKRTADNYVLLTAAQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPE 1544
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TAKRTADNYVLLTAAQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGCDACPGPE 540

Qy   1545 PSEVTRYQMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAP 1604
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 PSEVTRYQMCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAP 600

Qy   1605 VVDEEIVEECASFIPLEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAI 1664
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 VVDEEIVEECASFIPLEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATILAI 660

Qy   1665 IEYCCGLVTLPDNPFASCVFAPIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMM 1724
          ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db    661 IEYCCGLVTLPDNPFASCVFAPIAGITTPLPHKIKMFLSLFGGAIASKLTDARXALAFMM 720

Qy   1725 AGAAGTALGTWTSVGFVFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAA 1784
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AGAXGTALGTWTSVGFVFDMLGGYAGASSTACLTFKCLMGEWXTMDQLAGLVYSAFNPAA 780

Qy   1785 GVVGVLSACAMFALTTAGPDHWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTP 1844
          ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db    781 GVVGVLSACAMFALTTAGPDHWPNRLLTMLARSNTVCXEYFIATRDIRRKILGILEASTP 840

Qy   1845 WSVISACIRWLHTPEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYS 1904
          ||:||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db    841 WSXISACIRWLHTPEDDCGLIAWGLXIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYS 900

Qy   1905 CQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGS 1964
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGS 960

Qy   1965 ARPDPTDWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNVCFTQVPPTLRAAVAVDGVQV 2024
          |||||||||||||||||||||||:|:||||||||||||||||||||||||||||||||:|
Db    961 ARPDPTDWTSLVVNYGVRDYCKYEKLGDHIFVTAVSSPNVCFTQVPPTLRAAVAVADVRVQV 1020

Qy   2025 QCYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDALETNDCNSTNN 2084
          |:|||||||||||||||||||||||||||||||||||:||:||||||||:||||||:||
Db   1021 QXYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGGRMQLNLRDRLEANDCNSINN 1080
```

RESULT 5

US-08-424-550B-83

; Sequence 83, Application US/08424550B
; Publication No. US20020119447A1

US-08-424-550B-83

Query Match   48.0%; Score 7380; DB 8; Length 1422;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1392; Conservative 2; Mismatches 27; Indels 10; Gaps 2;

```
Qy  2085  TPSDEAAVSALVFKQELRRTNQLLEAISAGVDTTKLPAPS-IEEVVVRKRQFRARTGSLT  2143
                ||||||||||||||||||||||||||||||||||||  |:||||||||||||||||||||
Db  1081  TPSDEAAVSALVFKQELRRTNQLLEAISAGVDTTKLPAPSQIEEVVVRKRQFRARTGSLT  1140

Qy  2144  LPPPPRSVPGVSCCPESLQRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFFTAIGCAM  2203
                |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||
Db  1141  LPPPPRSVPGVSCCPESLQRSDPLEGPSXLPSSSPPVLQLAMPMPLLGAGECNPFFTAIGCAM  1200

Qy  2204  TETGGGPDDLPSYPPKKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPT  2263
                |||| :||:|||||||||||||||||  ||||||||||||||||||||||||:||||||||
Db  1201  TETKGXPXXLPSYPPKKEVSEWSDESWSTTTASSYVTGPPYPKIRGKDSTQSATARPT  1260

Qy  2264  KKKLGKSEFSCSMSYTWTDVISFKTASKVLSATRAITSGFLKQRSLVVVTEPRDAELRKQ  2323
                ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db  1261  KKKLGKSEFSCSMSYTWTDVISFKTASKVLSATRAITSGFLKQRSLVVVTEPRDAELRKQ  1320

Qy  2324  KVTINRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVR  2383
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  KVTINRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTD--  1378

Qy  2384  SGAARKAVLDLQKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLI  2434
                               |||||||||||||||||||||||||||||||||||
Db  1379  ------------VLDLQKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLI  1422
```

RESULT 6
us-10-467-000-1
; Sequence 1, Application US/10467000
; Publication No. US20040067486A1
; GENERAL INFORMATION:
;   APPLICANT: De Francesco, Raffaele
;   APPLICANT: Migliaccio, Giovanni
;   APPLICANT: Paonessa, Giacomo
;   TITLE OF INVENTION: HEPATITIS C VIRUS REPLICONS AND REPLICON
;   TITLE OF INVENTION: ENHANCED CELLS
;   FILE REFERENCE: ITR0003P
;   CURRENT APPLICATION NUMBER: US/10/467,000
;   CURRENT FILING DATE: 2003-07-21
;   PRIOR APPLICATION NUMBER: PCT/EP02/00526
;   PRIOR FILING DATE: 2002-01-16
;   PRIOR APPLICATION NUMBER: 60/263,479
;   PRIOR FILING DATE: 2001-01-23
;   NUMBER OF SEQ ID NOS: 13
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 3010
;   TYPE: PRT
;   ORGANISM: Con 1 HCV isolate nucleic acid
US-10-467-000-1

```
Query Match        21.2%;  Score 3262;  DB 15;  Length 3010;
Best Local Similarity   29.8%;  Pred. No. 1.3e-238;
Matches  970;  Conservative 452;  Mismatches 1183;  Indels 648;  Gaps 116;

Qy    8  TSPVPAPRTRKNKQTQASYPVSIK------------------TSVERGQRAKRKVQRDARPR   51
            |:  :|:|  |   | :                        ::|:
Db    3  TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR   59

Qy   52  NYKIAGIHDGLQTLAQAALP------------------------------AHGWGRQD   79
                :  |  ||   : ::                             || | |
Db   60  GRR--------QPIPKARQPEGRAWAQPGPWPLYGNEGLGWAGWLLSPRGSRPSWGPTD  111

Qy   80  PRHKSRNLGILLDYPLGWIGDVTTHTPLVGPLVAGAVRPVCQIVRLLEDGVNWATG---  136
             :  ||  : |||||   |  :  |||  |:|    |||:  :::|:| :   :
Db  112  PRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLP  170

Qy  137  WFGVHLFVVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC  195
             | :   :  :||  ||||| | |::   ||| : | :| |  : | || ||  |||
Db  171  GCSPSIFLALLSLCLTIPASAYEVRNV-SGVYHVTNDCSNASIVYEAADMIMHTPGCVPC  229

Qy  196  ADE------CWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVG  250
                      ||  |   :   |    :::| | |  | |  |::| :| :| | |::|
Db  230  VRENNSSRCWVALTPTLA-ARNASVPTTIRRHVDLLVGAAALCSAMYVGDLCGSVFLVA  288
```

```
Qy  1152 TKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYG 1211
         ||::| |:|:||||||||||| :: |:::|| |||| ::: ||| |||:||||
Db  1238 TKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGAPITYSTYG 1297

Qy  1212 MYLT-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPT 1269
         ::|| :|| ||:||||||||| :| ||||| ||| :|| ||||||||| ::
Db  1298 KFLADGGCSCSGAYDIIICDECHSTDSTTILGIGTVLDQAETAGARLVVLATATPGSVTV 1357

Qy  1270 PHANITEIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKCHCDELANELARKGITAVSY 1329
         || ||| :|:|:|| ||:|| | ||| | |||||| :: |||||| :| | |:||| :|
Db  1358 PHPNIEEVALSSTGEIPFYGKAIPIETIKGGRHLIFCHSKKKCDELAAKLSGLGLNAVAY 1417

Qy  1330 YRGCDISKIP-EGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLLDPTFTMGVRVC 1388
         |||  ::::| :|| :|||||||||:||||||||:|:|  | ||| ||||||||:|  ::|
Db  1418 YRGLDVSVIPTSGDVIVVATDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTV 1477

Qy  1389 GVSAIVKGQRRGRTGRGRMGIYRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAETSV 1448
         :| | : ||||||||||||||| || ||||||||:  | |||||||||:||::|||:|
Db  1478 PQDAVSRSQRQRGRTGRGRMGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQT 1537

Qy  1449 ILDTYRTQPGLPAIGANLDEWADLPS-MVNPEPSFVNTAKRTADNYVLLTAAQLQLCHQY 1507
         ||  | | |||| |:: |: |  :|: |:|:   :| || |||:| |:|| |:: :| :
Db  1538 RLRAYLNTPGLPVCQDHLEFWESVFTGLTHIDAHFLSQTKQAGDNFPYLVAYQATVCARA 1597

Qy  1508 GYAAPNDAPRWQG-ARLGKK----PCGVLWRLDGADACPGPEPSEVTRYQM-CFT---EVN 1559
         :: | ||||| | ||| || || :|| ||||:| ||| | :|
Db  1598 QAPPSWDQMWKCLIRLKPFLHGPTPLLYRL-GAVQNEVTTHPITKYIMACMSADLEVV 1656

Qy  1560 TSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI------VE 1613
         || | |||| |:|  :|| |:| |  | ::   ::| ||| ||:| |:
Db  1657 TS-TWVLVGGVLAAILAAYCLTTGSVVIVGR------IILSGKPAIIPDREVLYRFDEME 1709

Qy  1614 ECASFIP-LEAMVAAIDKLK-------STIT----TTSPPTLETALEKLNTFLGPHAATI 1661
         |||  || :| | |  ||:| |:   | || ||  ||||| || :||
Db  1710 ECASHLPYIEQGMQLAEQFKQKAIGLLQTATKQAEAAAP-VVESKWRTLEAFWAKHMWNF 1768

Qy  1662 LAIIEYCCGLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALA 1721
         :: | ||| |||| |:|   :|:| :|: |:| :|| :: ||| ::|| | |  ||:
Db  1769 ISGIQYLAGLSTLPGNPAIASLMAFTASITSPLITQHTLLFNILGGWVAAQLAPPSAASA 1828

Qy  1722 FMMAGAAGTALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYS 1778
         | | | ||:| :  |: |:   | |:| :| |:||   :|  :: |||| ::|| :| |
Db  1829 FVGAGIAGAAVG---SIGLGKVLVDILAGYGAGVAGALVAFKVMSGEMPSTEDLVNLLPA 1885

Qy  1779 AFNPAAGVGVLSACAMFALTTAGPD------HWPNRLLTMLARSNTVCNEYFIATRDIRRK 1834
         :: | | :||    | ||  |   |:      ::|| :  :: | | | :  |::
Db  1886 ILSPGALVVGVV--CAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAAR 1943

Qy  1835 ILGTILEASTPWSVISACIRWLHTPEDDCGLI---AWGLEIWQYVCNFFVICFNVLKAGV 1891
         | ::|  :  || :::: :||| :  :|| ::   :|  | |  |  |  :: |:
Db  1944 VTQILSSLT----ITQLLKRLHQWINEDCSTPCSGGSWLRDVWDWICT----VLTDFKTWL 1995

Qy  1892 QS--MVNIPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSN 1949
         ||  :: ||| || ||::|||| | | ::| :||||||::| | :||  ::|:|||||||
Db  1996 QSKLLPRLPGVPFFSCQRGYKGVWRGDGIMQTTCPCGAQITGHVKNGSMRIV-GPRTCSN 2054

Qy  1950 YWRGAVPVNARLCGGSARPDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFT 2007
         :| | | :||  ||:   | |: |:  |: |  |:::   ::|||::|||  ::||  |
Db  2055 TWHGTFPINAYTTGPCTPSPAPNYSRALWRVAAEEYVEVTRVGDFHYVTGMTTDNVKCPC 2114

Qy  2008 QVP------------------VDVGVQVCYLGEPKTPWTTS 2038
         |||                  ||:|:|||||| ::|
Db  2115 QVPAPEFFTEVDGVRLHRYAPACKFLLREEVTFLVGLNQYIVGSQLPC--EPE------ 2165

Qy  2039 ACCYGPDGKGKTVKL--PFRVDGHTPGVRMQLNLRDALETND-----------CNSTNN 2084
         :: : :: :::|  :    : | :  |  :: |  | |::||           |::: :
Db  2166 -----PDVAVLTSMLTDPSHITAETAKRRLARGSPPSLASSSASQLSAPSLKATCTTRHD 2220

Qy  2085 TPSDEAAVSALVFKQEL----RRTNQLL-----EAISAGVDTTKLPAPSIEEVVRK 2132
         :|:|: : ::| |::|:    || | ::     || :|||| :   ::|  ||:|
Db  2221 SPDADLIEANLLWRQEMGGNITRVESSENKVVILDSFEPLQAEBEDEREVSVPA--EILRRS 2278

Qy  2133 RQFRARTGSLTLPPPPRSVPGVSCP--------ESLQRSD----------PLBGPSNLPPSP 2176
```

```
Db  2279 RKP---------------------PRAMPIWARPDYNPELLESWKDPDYVPPVVHGCPLP-PAKAPPIP 2325
         |:|                     ||
Qy  2177 P------VLQLAMPMPLLGAGECNPF-----TAIGCAMTETGGGPDDLP 2214
         : |                   :::  ||
Db  2326 PPRRKRTVVLSESTVSSALELA--TKTFGGSESSAVDSGTATASPDQPSDDGDAGSDVE 2383

Qy  2215 SY---PP---KKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLG 2268
         ||   ||
Db  2384 SYSSMPPLEGEPGDPDLSDGSWSTVSEEAS 2413

Qy  2269 KSEFSCSMSYTWTD-VISFKTASKVLSATRAITSGFLKQRSLVYVTEPRDAELRKQKVTI 2327
         |::| |||||||| |:|:||||:||:|| |:|:   :|| ||:| ||  :|||||:||:
Db  2414 EDVVCCSMSYTWTGALITPCAAEETKLPINALSNSLLRHHNLVYATTSRSASLRQKKVTF 2473

Qy  2328 NRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGAA 2387
         :|: :  :::: || : :::||| | | | :: || ::|::||:      |||  :|:|
Db  2474 DRLQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKF-GYGAKDVRN-LS 2531

Qy  2388 RKAVLDLQK-----CVEAGEIPSHYRQTVIVPKEEVVKTPQKPTKKPPRLISYPHLEMRC 2443
         :||| : :|     || |:: | : | :  : |:|| :|::  :|| |||: : |:  |
Db  2532 SKAVNHIRSVWKDLLEDTETPI---DTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRV 2588

Qy  2444 VEKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMWSPD--AVGATCDTVCFDSTIT 2500
         :|||| :|| | |:| |||  |||  | |:||| || :| | |   | ||| |||||:|
Db  2589 CEKMALYDVVSTLPQAVMGSSYGFQYSPGQRVEFLVNAWKAKKCPMGFAYDTRCFDSTVT 2648

Qy  2501 PEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGMIAYDGREIGYRRCRSSGVYTTSS 2560
         ::|:  |  :|  ::||::|:| |: |:| | || |: |::|  : |||||:||:|||
Db  2649 ENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSC 2708

Qy  2561 SNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAP 2620
          |:|||:||  ||  : |  : ||::||||  |:  ||  :|: :| || | ||| |||
Db  2709 GNTLTCYLKAAAACRAAKLQDCTMLVCGDDLVICSSAGTQEDEASLRAFTEAMTRYSAP 2768

Qy  2621 QDCVPQPKYSLEELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWI 2680
          :| | |:|  |:||||||| :  :  :|| ||:||||| ||| : |:  ||:| :: ||
Db  2769 PGDPPKPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWL 2828

Qy  2681 GYLIHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPEDLPSIIAGVHGI 2740
         | :: | ||| ::: | ||:: : ::: ||  :   | | || | :|  ::: ::||
Db  2829 GNIIMYAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLHGL 2888

Qy  2741 EAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WH 2799
         :|||  :|:  ||  | :| ::  :||||| |::||:| | :::: :  :: :|| :|
Db  2889 SAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCGKYLFNWA 2948

Qy  2800 ATSR-----PLP----DLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVI 2850
          ::|     ||    ||    ::  |||    | |: ||                  :::
Db  2949 VRTKLKLTPIPAASQDLSSWFVAGYS-------------GGDIYHSLSRARPRW------F 2991

Qy  2851 VFALGLIAVGLAI 2863
         :: ||:  ||:||
Db  2992 MWCLLLLSVGVGI 3004
```

```
;  TYPE: PRT
;  ORGANISM: Artificial
;  FEATURE:
;  OTHER INFORMATION: HepC 1a consensus polyprotein
US-10-296-734-406
```

```
Query Match        21.0%;   Score 3234.5;  DB 15;  Length 3011;
Best Local Similarity  29.5%;  Pred. No. 1.6e-236;
Matches  959;  Conservative  437;  Mismatches 1214;  Indels 639;  Gaps 105;
```

```
Qy      8  TSPVPAPRTRKNKQTQASYPVSIK--------------------TSVERGQRAKRKVQRDARPR    51
           |:|:| :|:|:|   |                            :|:|:|| || |   || :|
Db      3  TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR    59

Qy     52  NYK----IAGIHDGLQTLAQAALP--------------------AHGWGRQDPRHKSRN    86
           :|| |    |  | | :| :   |                         | |:|:|:| |||
Db     60  GRRQPIPKARRPEG-RTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRN   118

Qy     87  LGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG--WFGVHLF   143
           | :||  | :: | :| :  | || |  || | ||| :|| : |     |    :||| |
Db    119  LGKVIDTLTCGFADLMGYIPLVGAPLGGA--ARALAHGVRVLEDGVNYATGNLPGCSFSIF   177

Qy    144  VVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC------AD   197
           ::| ||  :|  | | | ||  :  | |||    |:| ||  | : ||  :|         |
Db    178  LLALLSCLTVPASAYQVRN-STGLYHVTNDCPNSSIVYEAADAILHTPGCVPCVREGNAS   236

Qy    198  ECWVPANPYI-SHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV--   254
           || |||  |   | ::| |  |:   | |:| | | |  | || : | :| || |  |
Db    237  RCWVAMTPTVATRDGKLPATQ--LRRHIDLLVGSATLCSALYGDLCGSVFLVGQLFTFS   294

Qy    255  --RHWLIHIDLNETGTCVLEVPTGIDPGFLGFIGW----MAGKVEAVIFLTKLASQVPYAI   309
             ||   || |: || ||  |  | |  | |: |       |  |  |  |||   |  :|
Db    295  PRRHW------TTQGCNCSIYPGHITGHR--MAWDMMMNWSPTAALVMAQLL-RIPQAI   344

Qy    310  ATMFSSVHYLAVAGALIYYASRGKWYQLLLALMLY------IEATSGNPIRVPTGC----   358
           |: ||:|| |||| | |:||:| |||:||| | :|      :| |:| | ||  |
Db    345  LDMLAGAHWGVLAGIAYFSMVGNWAKVLVVLLLFAGVDAETHVTGGNAGRTTSGLVSLLT   404

Qy    359  SIAEPCS------------------------------------------------------   365
           :|   ||
Db    405  PGAKQNIQLINTNGSWHINSTALNCNESLNTGWLAGLFYQHKFNSSGCPERLASCRRLTD   464

Qy    366  PLMI--------PCPCHSYLSENV-SEVICYSPKWTRPITLEYNN-                401
           || :        | || || :||  || ||||: :  |  |:|::
Db    465  FDQGWGPISYANGSGPDRPYCWHYPPKPCGIVPAKSVCGPVVCFTPS---PVVVGTTDR   521

Qy    402  --SISWYPYTIPGARGCMVKFKQNT-----WGCCRIRNVPSY-----CTMG---         440
             :|| | || ||  ||  |||::|     ||| :|| ||::     |  |
Db    522  SGAPTYSW------GANDTDVFVLNNTRPPLGNWFGCTWMNSTGFTKVCGAPPCVIGGAG   575

Qy    441  --------TDAVWNDTRNTYEACGVTPWLT----------TAWHNGSALKLAILQ--YPGS   481
                   | | |:|:| ||| ||| ||             | |: :|:||    |
Db    576  NNTLHCPTDCFRKHPEATYSRCGSGPWITPRCLVDYPYRLWHYPCTINYTIFKVRMYVGG   635

Qy    482  KE--MFKPHNWMSG-HLYFEGSDTPIVIFYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQV   538
           |    |||:| ||  ||| | ||  |  |   |  :  | | || || |: : || || |
Db    636  VEHRLBAACNWTRGERCDLEDRDR----SELSPLLLSTTQWQVLPCS--------         678

Qy    539  PQGFYSDVDKLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILGLCGSKYLIL   598
           | |||| | | | | |  |  |||| | :| | || :||  |||| |:|| |  | ::
Db    679  ----FTTLPALSTGLIHLHQNIVDVQYLYGVGSSIASWAIKWEYVVLLFLLLADARVCSC   734

Qy    599  AYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYA   658
           | :| ::|| : | ||     ||  |  | |    ||   | |:  ::|| | :|| |
Db    735  LWMMLL------------------ISQAEAALENLVI-----LNAA               757

Qy    659  ALLGFVPMAAGLP--LTFFVAAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALLVGPWP   716
           ||:|| |  |||   | :|  ::||  | |: |:||: | |||| | |   | ||||
Db    758  SLAG----THGLVSFLVFCFA-------WYLK--------GRWVPGAVTALYGMWP   795

Qy    717  LVALLTLLHLVTPASAFDTEI---IGGLTIPPVVALVVMSRFGFPAHLLPRCALVNSYLW   773
           || |:|| |:|   | |         |::|| |||  : ||  |  | |  || | :| |
Db    796  L---LLLLLALPQRAYALDTEVAASCGGVVL---VGLMALTLSPYYKRYISWCLWWLQYFL   850

Qy    774  QRWEN-----WFWNVTLRPERFFLVCFPGATYDALVTFCVCHVALLC-LTSSAASFFGT   828
           |||:|     || |:|| |  | ||:| || | ||| |:| ||:      ||| |:|
Db    851  TRVEAQLHVWVPPLNVRGGRDAVILL--------------MCVVHPTLVPDITKLLLAVFGP   898

Qy    829  DSRVRAHRMLV-------------RLGKCHAWYSHYVLKFPLLVFGENGVFPFYKH   870
           ||||| | |:               | | |:  :|:|  || : | | | |||:|:
Db    899  LWIIQASLLKVPYFVRVQGLLRICALARKMIG------GHYVQMAIIKLGALTGTYVYNH   952

Qy    871  LH--GDVLPNDFASKLPLQEPFFPPEGKARVYRNEGRRLACGDTVDGLPVVARLGDLVFA   928
           |     |:||:| ||| | |    | |     : |:| :|   |||  ||| |:|| |
Db    953  LTPLRDWAHNGLRDLAVAVEPVVFSQMETKLITWGADTAACGDIINGLPVSARRGREILL  1012

Qy    929  GLA--MPPDGWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSYMGF   986
           | |    || |   | :  ||   || |   |  :|| :  ||  | ||| |: |:   |
Db   1013  GPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQISVSTAAQTFLAT  1072

Qy    987  VCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDI--YQPPCGAGSLTRCSCGETKG  1044
           | || | | | |  | | | |:| :|   || :|:    ||     |: | ||  |
Db   1073  CINGVCWTVYHGAGTRTIASPKGPV--IQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDL  1130

Qy   1045  YLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPLLCSSGHVIGMFTAA--RNSG  1101
           ||| | | |  |   |   | |||| |   |  :| | |  :| | |  |:|     |:|
Db   1131  YLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVA  1190

Qy   1102  GSVSQIRVRPLVCAGYHPQTAHATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQE  1161
           | |  :|| |   ||  |   |: |  || || || : || : ||||| || ||:||
Db   1191  KAVDFIPVENLETTMRSPVFTDNS---SPPAVPQSFQVAHLHAPTGSGKSTKVPAAYAAQ  1247

Qy   1162  KYEVLVLNPSVATTASMPKYMHATGVNPNCYFNGKCTNTGASLTYSTYGMYLT-GACSR  1220
           | || |||||||  |:  |:: : |   | ||| || |   |: |||| ||||
Db   1248  GYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFLADGGCSG  1307

Qy   1221  N-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANITEIQL  1279
           |   |::|||||:  | :|  | |  |  |  | | |  |||||| : |  | :|:|
Db   1308  GAYDIIICDBCHSTDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVAL  1367

Qy   1280  TDEGTIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELARKGITAVSYYRGCDISKIP  1339
           | :| ||| |||::||:||| |||  | :|:||::| | | ||| ||   |:|| |  :
Db   1368  STTGEIPFYGKAIPLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIP  1427

Qy   1340  -EGDCVVVATDALCTGYTGDFDSVVDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQR  1398
            || |||||||||||||||||||||||:|  |||||| |||||||  |||||||||:||:
Db   1428  TSGDVVVVATDALMTGYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTLPQDAVSRTQR  1487

Qy   1399  RGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPG  1458
           |||||||||||  |  | |   |  |   |  |    |  | : |  |  | |  | || |
Db   1488  RGRTGRKGPGIYRFVAPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPG  1547

Qy   1459  LPAIGANLDBWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLCHQVGYAAPNDAPR  1517
           || | | || |  :    | |:||| |  | | || |||  ||| :| |  | | | ||
Db   1548  LPVCQDHLBFWEGVFTGLTHIDAHFLSQTKQSGENFPYLVAYQATVCARAQAPPSWDQM  1607

Qy   1518  WQG-ARLGKK---PCGVLWRLDGADACPGPEPSEVTRYQM-CFT--EVNTSGTAALAVG  1569
           | |     |    ||  | |     |   | |  : | ::  |    || | | :|| |
Db   1608  WKCLIRLKPTLHGPTPLLYRL-GAVQNEVTLTHPVTKYIMTCMSADLEVVTS-TWVLVGG  1665

Qy   1570  VGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI----VEECASFIPL--   1621
           :| :|  | | |   | :|||:  :| |    | | ::         |: ||| :|
Db   1666  VLAALAAYCLSTGCVVIVGR------IVLSGKPAIIPDREVLYREFDEMEECSQHLPYIE  1719

Qy   1622  EAMVAA----IDKLKSTITTS--------PFTLETALEKLNTFLGPHAATTLAIIEYCCGL  1671
           :|::    :||: |  :: |           |   | |:|  : |    |::   |:||
Db   1720  QGMMLAEQFKQKALGLLQTASRQABVIAP-AVQTNWQKLEVFWAKHMWNFISGIQYLAGL  1778

Qy   1672  VTLPDNPFASCVFAFIAGITTLPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAGTA  1731
           | ||:|| |   | ||  :| : |  |   |  | || :|      || | | : |  | |
Db   1779  STLPGNPAIASLMAFTAAVTSPLTTSQTLLFNILGGWVAQLAAQAAAPGAATAFVGAGLAGAA  1838

Qy   1732  LGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVVG  1788
           ||  | ||    : :::| ||  :|   | | |   | :| |   | | :: : | | | |
Db   1839  IG---SVGLGKVLVDILAGYGAVGAGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVVG  1895
```

Qy   1789 VLSACAMFALTTAGPD-----HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEASTP 1844
            :||: || :|| || ::: || :||||||||||||||||||||
Db   2889 LSAFSLHSYSPGEINRVAACLRKLGVPPPLRAWRHRARSVRARLLARGGRAAICGKYLFNW 2948

Qy   1896 VV--CAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAARVTAILSSLTV 1953
            ::   ::||||| :::   : ::
Db   2799 HATSR----PLPDLDKTSVARYTFNYCDVVSPEGDVFITPQRRLQKPLVKYLAVIVFAL 2854

Qy   1845 WSVISACIRWLHTPTEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFYS 1904
            :: ::|| | :|| :| || ::   :
Db   2949 AVRTKLKLTPIAAAGRLDLSGWFTAGYSG-----GDIYHSVSHARPRW-------FWFCL 2996

Qy   1954 TQLLRLHQWISSECTTPCS-GSWLRDIWDWICEVLSDFKTWLKA--KLMPQLPGIPFVS 2010
            :|:: | | :|: :|| :   :|

Db   2855 GLIAVGLAI 2863
            :|: ::|:
Db   2997 LLLAAGVGI 3005

Qy   1905 CQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCGS 1964
Qy   2011 CQRGYKGVWRGDGIMHTRCHCGABITGHVKNGTMRIV-GPRTCRNMWSGTFPINAYTTGP 2069

Qy   1965 ARPDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFTQVP----------- 2010
Db   2070 CTPLPAPNYTFALMRVSAEEYVEIRRVGDFHYVTGMTTDNLKCPCQVPSPEFFTELDGVR 2129

Qy   2011 -------PTLRAAVAVD-----------GVQVQCYLGEPKTPWTTSACCYGPDGKGKTVKL 2053
Db   2130 LHRFAPPCKLLREEVSFRVGLHEYPVGSQLPC----EPE------------PDVAVLTSML 2175

Qy   2054 --PPRVDGHTPGVRMQLNLRDALETND----------CNSTNNTPSDEAAVSALVFKQ 2099
Db   2176 TDPSHITAEAAGRRLARGSPPSMASSASQLSAPSLKATCTANHDSPDAELIEANLLWRQ 2235

Qy   2100 EL------RRTNQLL-----EAISAGVDTTKLPAPSIEVVVRKRQFRARTGSLTLP-- 2145
Db   2236 EMGGNITRVESENKVVILDSFPDLVAEEDEREISVPA--EILRKSRRFAQALPVWARPDY 2293

Qy   2146 --------------PPPRSVP----------- 2152
                             ||||||
Db   2294 NPPLVETWKKPDYEPPVVHGCPLPPPRSPPVPPPRRKKRTVVLTESTLSTALAELATKSFG 2353

Qy   2153 ----GVSCPESLQRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGBCNPFTAIGCAMTETG 2207
Db   2354 SSSTSGITGDNTTTSSEP--APSGGCPPDSDAESYSSMMPLEG----------------EP 2395

Qy   2208 GGPDDLPSYPKKEVSYPKKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKL 2267
            ||:|||
Db   2396 GDPD------------------------------LSDGSWSTVSSEAG- 2413

Qy   2268 GKSEFSCSMSYTWTDVISFKTASKVLS-ATRAITSGFLKQRSLVVTEPRDAELRKQKVT 2326
            ||||||| :: ::
Db   2414 TEDVVCCSMSYSWTGALVTPCAAEEQKLPINALSNSLLRHHNLVYSTTSRSACQRQKKVT 2473

Qy   2327 INRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSGA 2386
Db   2474 FDRLQVLDSHYQDVLKEVKAAASKYKANLLSVEEACSLTPPHSAXSKF-GYGAKDVRC-H 2531

Qy   2387 ARKAVLDL----QKCVEAGEIPSHYRQTTVTVPKEEVFVKTPQKPTKKPPRLISYPHLEMR 2442
Db   2532 ARKAVAHINSVWKDLLEDSVTPI---DTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVR 2588

Qy   2443 CVEKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMW--SPDAVGATCDTVCFDSTI 2499
Db   2589 VCEKMALYDVVSKLPLAVMGSSYGFYSPGQRVEFLVQAWKSKKTPMGFSYDTRCFDSTV 2648

Qy   2500 TPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTS 2559
Db   2649 TESDIRTEEAIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGVLTTS 2708

Qy   2560 SSNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGA 2619
Db   2709 CGNTLTCYIKARAACRAAGLQDCTMLVCGDDLVVICESAGVQEDAASLRAFTEAMTRYSA 2768

Qy   2620 PQDCVPQPKYISLEELTSCSSNVTSGITKSGKPYYFLTRDRPIPLGRCSAEGLGYNPSAAW 2679
Db   2769 PPGDPPQPEYDLELITSCSSNVSVAHDGAGKRVYYLTRDPTTPLARAAWETARHTPVNSW 2828

Qy   2680 IGYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVHG 2739
Db   2829 LGNIIMFAPTLWARMLMTHFFSVLIARDLEQALDCEIYGACYSIEPLDLPPIIQRLHG 2888

Qy   2740 IEAFSVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-W 2798
            :: | | ::: |||
Db   2740 -YSPKWTRPITLEYNNSISWYPY-----------TIPGARGC------ 2740

RESULT 8
US-10-333-449A-34
; Sequence 34, Application US/10333449A
; Publication No. US20040137424A1
; GENERAL INFORMATION:
; APPLICANT: Tan, Yin Hwee
; APPLICANT: Lim, Siew Pheng
; APPLICANT: Lim, Seng Gee
; APPLICANT: Hong, Wan Jin
; TITLE OF INVENTION: NUCLEIC ACIDS AND METHODS FOR DETECTING VIRAL INFECTION,
; TITLE OF INVENTION: UNCOVERING ANTI-VIRAL DRUG CANDIDATES AND DETERMINING DRUG
; TITLE OF INVENTION: RESISTANCE OF VIRAL ISOLATES
; FILE REFERENCE: 01/22137
; CURRENT APPLICATION NUMBER: US/10/333,449A
; CURRENT FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 34
; LENGTH: 3010
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-333-449A-34

  Query Match   21.0%; Score 3232; DB 16; Length 3010;
  Best Local Similarity   29.9%; Pred. No. 2.6e-236;
  Matches 964; Conservative 451; Mismatches 1222; Indels 586; Gaps 116;

Qy   9 SPVPAPRTRKNKQTQASYPVSIK--------------TSVERGQRAKRKVQRDARPRN 52
            |:   :|:
Db   2 STIPKPQ-RKTKRNTYRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60

Qy   53 YKIAGIHDGLQTLAQAALP-----------------------------AHGWGRQDP 80
            :   :|| ||:
Db   61 RR--------QPIPKARRPEGWAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPIDP 112

Qy   81 RHKSRNLGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG--W 137
            :||:: |::| :| |: ::| ||| : |||:
Db   113 RRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLPG 171

Qy   138 FGVHLFVVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVICA 196
            |||:| |||:   ::| : |:| :| :| | ||::|
Db   172 CSFSIFLALLSCLTIPASAYEVRNA-SGVVHVTNDCSNSSIVYEAADIIMHTPGCVPCV 230

Qy   197 DE---------CWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGD 251
Db   231 RENNSSRCWVALTPTLA-ARNLSVPTTIRRHVDLLVGAAAFCSAMYVGDLCGSVLLVSQ 289

Qy   252 WLV-----RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGWMAGKVEAVIFLTKLASQVPY 307
Db   290 LFTLSPRQHETVQ-DCN---CSL-YPGHVTGHRMAWDMMNWSPTTALVLSQIL-RIPQ 342

Qy   308 AIATMFSSVHYLAVGALIYYASRGKWYQLLLALMLYI------------EATSGNPI-- 352
Db   343 TIVDMVAGAHWGVLAGIAYYSMVGNWAKVLVVMLLFAGVDGHTQVMGGSQASTINTLTGI 402

Qy   353 ------------------------------RVPTGCSIAEFCSPLMIPCPCHSYLSENVSEVIC----- 386
Db   403 FSPGAKQIQLINTNGSWHINRTALNCNDSLNTGFLAALFYTHSFNSSGCLERMASCRPI 462

Qy   387 --YSPKWTRPITLEYNNSISWYPY------------TIPGARGC--- 416

```
Top block:

1402 Qy  TGRGRAGIYYVVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTILDYTROPGLPA   1461
1491 Db  TGRGRGGIYRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAETSVRLRAYLNTPGLPV   1550

1462 Qy  IGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLQLCHQYGYAAPNDAPRWQG   1520
1551 Db  CQDHLEFWESVFTGLTHIDAHFLSQTKQAGDNFPYLVAYQATVCARAQAPPSWDQMWKS    1610

1521 Qy  -ARLGKK---PCGVLWRLDGADACPGPEPSE-------VTRYQM-CFT---EVNTSGTAAL  1566
1611 Db  LIRLKPTLHGPTPLLYRL------GTVQSEITLTHPVTKYIMACMSADLEVVTS-TWVL    1662

1567 Qy  AVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI------VEECASFIP    1620
1663 Db  VSGVLAALAAYCLTTGSVVIVGR------IILSGKPAVIPDREVLYREPDEMEECASHLP  1716

1621 Qy  -LEAMVAAIDKLK--------STIT----TTSPFTLETALEKLNTFLGPHAATILAIIEYC  1668
1717 Db  YIEOQMQLAEQFKQKALGLLQTATKQAEAAAP-VVESKWRALEAFWAKHMWNFISGIQYL   1775

1669 Qy  CGLVTLPDNPFAASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAA 1728
1776 Db  AGLSTLHGNPAIASLMAFFTASITSPLTTQHTLLFNILGGWVAAQLAPPSAASAFVGAGIA 1835

1729 Qy  GTALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAG  1785
1836 Db  GAAVG---SIGLGKVLVDVLAGYGAVGAGLVAFKVMSGEVPSTEDLVNLLPAVLSPGAL    1892

1786 Qy  VVGVLSACAMFALTTAGPD----HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEA  1841
1893 Db  VVGVV---CAAILRRHVGPGEGAVQWMNRLLIAFASRGNHVSPTHVPESDAAARVTQILSS 1950

1842 Qy  STPWSVISACIRWLHTPTEDDCGLI---AWGLEIWQYVCNFFVICFNVLKAGVQS--MVN   1896
1951 Db  LT----VTQLLRRLHQWINEDCSTPCSGSWLRDVWDWICT----VLTDFKTWLQSKLLPR  2002

1897 Qy  IPGCPFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVP  1956
2003 Db  LPGVPFFSCCORGYRGVWRGDGIMQTTCPCGAQISGHVKNGSMRIV-GPRSCSNTWHGTFP 2061

1957 Qy  VNARLCGSARPDPT-DWTSLVVNVGNVRDYCKYEKMGDHIPVTAVSSPNV-CFTQVP----  2010
2062 Db  INAYTTGPCTPAPAPNYSRALWRVAAEEYVEVTRVGDFHVVTGMTTDNVKCPCQVPAPEF  2121

2011 Qy  ----------------PTLRAAVAVD----------GVQVQCYLGEPKTPWTTSACCYGPD  2045
2122 Db  FTEVDGVRLHRYAPACKPLLREEVTFQVGLNQYLVGSQLPC---EPE----------PD   2167

2046 Qy  GKGKTVKL--PFRVDGHTPGVRMQLNLRDALETND---------CNSTNNTPSDEAA     2091
2168 Db  VAVLTSMLTDPSHITAETAKRLLDRGSPPSLASSSASQLSAPSLKATCYTRHDSPDAGLI  2227

2092 Qy  VSALVFKQEL-------RRTNQLL-----EAISAGVDTTKLPAPSIEEVVRKRQFRART    2139
2228 Db  EANLLWRQEMGGNITRVESENKVVILDSFEPLRAEEDEREVSVPA--EILRKSRKFPRAM  2285

2140 Qy  GSLTLP--PP-------PRSVPGV--SCPESLQRSDPLEGPSNLPPSPP-------VLQ   2180
2286 Db  PIWARPDYNPLLESWKNPDYVPPVVHGCP-----LPPIKG---PPIPPPRRKRTVVLT    2336

2181 Qy  LAMPMPLLGAGECNPFTAIGCAMTETG--GGPDDLPSYPPKKEVSEWSDESWSTATTASS  2238
2337 Db  ESTVSSALAELATKTFGSSGSSAVDGSTASAPPDOPS-----------DNGDAGSDAES   2384

2239 Qy  YVTGPPYPKIRGK--DSTQSAPAKRPTKKKLGKSEFSCCSMSYTWTD-VISFKTASKVLSA 2295
2385 Db  YSSMPP---LEGEPGDPDLSDGSWSTVSBEASEDVVCCSMSYTWTGALITPCAAEESKLP  2441

2296 Qy  TRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPPSYHKQVRLAKEKASKVVGVM  2355
2442 Db  INALSNSLLRHHNMVYATTSRSASQRQKKVTFDRLQVLDDHYRDVLKEMCAKASTVKAKL  2501
```

```
Bottom block:

463  Db  DKFPDQGW-GPITYDEGPDLDQRPYCWHYAPRSCGIVPASQVCGPVYCFTPSPVVGTDR    521
417  Qy  ------------MVKFKNT------WGCCRIRNVPSY------------CTMG--------  440
522  Db  TGAPTYRWGENETDVLILNNTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGGAGNNTLVC   581

441  Qy  -TDAVWNDTRNTYEACGVTPWLT--------TAWHNGSALKLAILQ--YPGSKE-MF      485
582  Db  PTDCFRKHPEATYTKCGSGSPWLTPRCMVDYPYRPWHYPCTVNFTIFKVRMYVCGVEHRLN 641

486  Qy  KPHNWMSG-HLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVPQGFYS   544
642  Db  AARNWTRGERCDLEDRDR------SELSPLLLSTTEWQILPCS-----------FT       680

545  Qy  DVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILGLCGSKYLILAYLCYL    604
681  Db  TLPALSTGLIHLHQNIVDVQYLY----------GIGSVVVSLVIKW---EYVLLSFF---  724

605  Qy  SLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAALLGFV   664
725  Db  ----------------LLADARVCACLWMMLLIAQAEAALENLVVLNAA              757

665  Qy  PMAA--GLPLTFFVAAAAQPDYDWWVR-LLVAGLVLWAGRNRGHRIALLVGPWPLVALL    721
758  Db  SVAGAHGI-LTFLVFFCAA-----WYIKGRLVPGAAY-----------AFYGVWPL--LL  798

722  Qy  TLLHLVTPASAFDTEI---IGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQRWEN   778
799  Db  LLLALPPRAYAMDREMAASCGGAVFIGLVLLTLSPHYKVFLARL----IW------      844

779  Qy  WFWNVTLRPERRFPLVLVCFP----GATYDALVTFCVCH----------VALLCJTSSAASF 825
845  Db  WLQYFTTRAEAILHVWV--PPLNVRGGRDAVILLTCAVHPDLIFDITKLLLAVLGPLMVF   902

826  Qy  FGTDSR-----VRAHRML--VRLGKCHAWYSHYVLKFFLLVFGENGVFFYKHL------H  872
903  Db  LAGITRVPYFVRAQGLIRACALARKVAG--GHYIQMALMKLAALTGTYLYDHLTPLRDWAH  961

873  Qy  GDVLPNDFASKLPLQEPFFPFEGKARVYRNEGRRLACGDTVDGLPVVARLGDLVFAGLA-  931
962  Db  AGL--RDLAVAV---EPVVFSDMETKIITWGADTAACGDIILGLPVSARRGREILLGPAD  1016

932  Qy  -MPPDGWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSYMGFVCDN  990
1017 Db  SLEGQGWRLLAPITAYSQQTRGLLGCIITSLTGRDKNQVEGEVQVVSTATQSFLATCVNG  1076

991  Qy  VLYTAHHGSKGRRLAHPTGSIHPITVDAAN-DQDI--YQPPCCAGSLTRCSCGETKGYLV  1047
1077 Db  VCWTVFHGAGSKTLAGPKG---PITQMYTNVDQDLVGWQAPPGGARSLITPCTCGSSDLYIV 1133

1048 Qy  TRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCGSSGHVIGMFTAA---RNSGGSV 1104
1134 Db  TRHADVIPVRRRGDNRGSLLSPRPVSYLKGSSSGGPLLCPSGHAVGIFRAAVCTRGVAKAV 1193

1105 Qy  SQIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILLAPTGSGKSTKLPLSYMQBKYE  1164
1194 Db  DFIPVESMETTMRSPVFTDNS---SPPAVPQTFQVAHLHAPTGSGKSTKVPAAYAAQGYK  1250

1165 Qy  VLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYLT-GACSRN-Y  1222
1251 Db  VLVLNPSVAATLGFGTYMSKAYGVDPNIRTGVRTITTGAPITYSTYGKFPLADGGCSGGAY 1310

1223 Qy  DVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTHANITEIQLUDE   1282
1311 Db  DIIMCDECHSTDSTTVLGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEIALSNT  1370

1283 Qy  GTIPFHGKKIKEENLKKGRHLIFEATKGHCDELANELARKGITAVSYYRGCDISKIP-EG  1341
1371 Db  GEIPFYGKAIPIETIKGGRHLIFCHSKKKCDELAAKLSGLGLNAVAYYRGLDVSVIPTSG  1430

1342 Qy  DCVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGR   1401
1431 Db  DVVVATDALMTGFTGDSDSVIDCNTCVTQTVDFSLDPTFTIETTVPQDAVSRSQRRGR    1490
```

```
Qy    60 GRRQPIPKARRPEG--RTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRN 118
Db    87 LGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLF 143
Qy   119 LGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLPGCSFSIF 177
Db   144 VVCLLS-LACPCSGARVTDPPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC----AD 197
Qy   178 LLALLSCLTVPASAYQVRN-SSGLYHVTNDCPNSSIVYEAADAILHTPGCVPCVREGNAS 236
Db   198 ECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV--- 254
Qy   237 RCWVAVTPTVATRDGKLPTTQ-LRRHIDLLVGSATLCSALYVGDLCGSVFLVGQLFTFSP 295
Db   255 -RHWLIHIDLNETGTCYLEVPTGIDPGFLGPIGW--MAGKVEAVIFLTKLASQVPYAIA 310
Qy   296 RRHW----TTQDCNCSIYPGHITGHR-MAWDMMMNWSPTAALVVAQLL-RIPQAIM 345
Db   311 TMFSSVHYLAVGALIYYASRGKWYQLLLALMLY-IEA 346
Qy   346 DMIAGAHWGVLAGIAYFSMVGNWAKVLVLLFAGVDAETHVTGGSAGHTTAGLVGLLTP 405
Db   347 TSGNPIRVPTGCSIAEF 363
Qy   406 GAKQNIQLINTNGSWHINSTALNCNESLNTGWLAGLFYQHKFNSSGCPERLASCRRLTDF 465
Db   364 C---SPLMI PCPCHSYLSENV-SEVICYSPKWTRPITLEYNN- 401
Qy   466 AQGWGPISYANGSGLDERPYCWHYPPRPCGIVPAKSVCGPVYCFTPS---PVVVGTTDRS 522
Db   402 ---SISWYPYTIPGARGCMVKFKNNT WGCCRIRNVPSY CTMG--- 440
Qy   523 GAPTYSW-------GANDTDVFVLNNTRPPLGNWFGCTWMNSTGFTKVCGAPPCVIGVGVGN 576
Db   441 TDAVWNDTRNTYEACGVTPWLT TAWNGSALKLAIIQ---YPGSK 482
Qy   577 NTLLCPTDCFRKHPEATYSRCGSGPWITPRCMVDYPYRLWHYPCTINYTIFKVRMYVGGV 636
Db   483 E--MFKPHNVMSG--HLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVP 539
Db   637 EHRLEAACNWTRGERCDLEDRDR SELSPLLLSTTQWQVLPCS 678
Qy   540 QGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILA 599
Db   679 FTTLPALSTGLIHLHQNIVDVQVLYGVGSSIASWAIKWEYVVLLFLLLADARVCSCL 735
Qy   600 YLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPPALIFPICCYLRCRLRYAA 659
Db   736 WMMLL ISQAEAALENLVI------ LNAAS 758
Qy   660 LLGFVPMAAGLP--LTFFVAAAAQPDYDWVRLLVAGLVLWAGRNRGHRIALLVGPWPL 717
Db   759 LAG----THGLVSFLVFFCFA-----WYLK GRWVPGAVYAFYGMWPL 796
Qy   718 VALLTLLHLVTPASAFDTEI---IGGLTIPPVALVVMSRFGFFAHLLPRCALVNSYLWQ 774
Db   797 --LLLLLALPQORAYALDTEVAASCGGVVL--VGLMALTLSPYYKRYISWCMWLQYFLT 851
Qy   775 RWEN----WFWNVTLRPERFFLVLVCFPGATYDALVTFCVCHVALLC-LTSSAASFGT- 828
Db   852 RVEAAQLHVWVPPLNVRGGRDAVILL MCVVHPTLVFDITKLLLAIFGPL 899
Qy   829 DSRVRAHRMLVRLGKCHAWYSHYVLKFFLLVFGENGVFFYKLH--GD 874
Db   900 WILQASLLKVPYFVRVQGLLRICALARKIAG-GHYVQMAIIKLGALTGTYVVNHLTPLRD 958
Qy   875 VLPNDFASKLPLQEP--FFPPEGKARVYRNEGRRLACGDTVDGLPVVARLGDIVPAGLA- 931
Db   959 WAHNGLRDLAVAVEPVVFSRMETKLLITWGAD--TAACGDIINGLPVSARRGQEILLGPAD 1016
Qy   932 -MPPDGWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSYMGFVCDN 990
Db  1017 GMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTATQTFLATCING 1076
```

```
Qy  2356 WDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDLQK----CVEAGEIPSHYRQTV 2411
Db  2502 LSVEEACKLTPPHSAKSKF-GYGAKDVRN-LSSKAVNHIRSVWKDLLEDTETPI---DTT 2556
Qy  2412 IVPKEEVFVKTPQKPTKKPPRLISYPHLEMRCVEKMYGQVAPDVVKAVMGDAYGF-VDP 2470
Db  2557 IMAKNEVFCVQPEKGGRGKPARLIVFPDLGVRVCEKMALYVDVVSTLPQAVMGSSYGFQYSP 2616
Qy  2471 RTRVKRLLSMW------SPDAVGATCDTVCFDSTITPEDIMVETDIYSAAKLSDQHRAGIHT 2526
Db  2617 GQRVEFLVNAWKSKKSP--MGFAYDTRCFDSTVTESDIRVEESIYQCCDLVPEARQAIRS 2674
Qy  2527 IARQLYAGGPMIAYDGREIGYRRCRSSGVVTTSSSNSLTCWLKVNAAAEQAGMKKNPRFLI 2586
Db  2675 LTERLYVGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKASACRAAKLQDCTMLV 2734
Qy  2587 CGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSNVTSGIT 2646
Db  2735 CGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHD 2794
Qy  2647 KSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHYPCLWVSRVLAVHFMEQMLF 2706
Db  2795 ASGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTLWARMILMTHFFSILLA 2854
Qy  2707 EDKLPETVTFPDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSLTDMTMP 2766
Db  2855 QEQLEKALDCQIYGAYYSIEPLDLPQIIERLHGLSAFSLHSYSPGEINRVASCLRKLGVP 2914
Qy  2767 PLRAWRKKARAVLASAKRRGGAHAKLARFLL-WHATSR----PLPDLDKTSVARYTTFNY 2821
Db  2915 PLRAWRHRARSVRAKLLSQGGRAATCGKYLFNWAVKTKLKLTPIPAASQLDLSGWFVAGY 2974
Qy  2822 CDVVSPEGDVFITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS 2864
Db  2975 -----GGGDIYHSLLSRARPRW-------FMLCLPLLSVGVIN 3005
```

RESULT 9
US-09-742-659-4
; Sequence 4, Application US/09742659
; Patent No. US2001003401A1
; GENERAL INFORMATION:
; APPLICANT: Hong, Zhi
; APPLICANT: Butkiewicz, Nancy J.
; APPLICANT: Zhong, Weidong
; APPLICANT: Ingravallo, Paul
; APPLICANT: Wright-Minogue, Jacquelyn
; APPLICANT: Lau, Johnson Y.
; APPLICANT: Lemon, Stanley M.
; TITLE OF INVENTION: Chimeric HCV/GBV-B viruses
; FILE REFERENCE: ID01116
; CURRENT APPLICATION NUMBER: US/09/742,659
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/171,469
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 4
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-742-659-4

Query Match 20.9%; Score 3219.5; DB 9; Length 3011;
Best Local Similarity 29.4%; Pred. No. 2.3e-235;
Matches 957; Conservative 444; Mismatches 1207; Indels 643; Gaps 108;

```
Qy     8 TSPVPAPRTRKNKQTQASYPVSIK-----------TSVERGQRAKRKVQRDARPR 51
Db     3 TNPKPQRKTKRNTNR---PQDVKFPGGGQIVGGVLLPRRGPRLGVRATRKTSERSQPR 59
Qy    52 NYK----IAGIHDGLQTLAQAALP-------------AHGWGRQDPRHKSRN 86
```

```
Qy    991  VLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDI--YQPPCGAGSLTRCSCGETKGYLVT 1048
                ::||:    ||||||||   :::  ||: |  |          :    :||||:|||||
Db   1077  VCWTVYHGAGTRTIASPKGPV--IQMYINVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVT 1134

Qy   1049  RLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAA---RNSGGSVS 1105
           |  ::   :   ||  ::|::  |  |||||||| ||  ||::| | ||              :
Db   1135  RHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAVCTRGVAKAVD 1194

Qy   1106  QIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILLAPTGSGKSTKLPLSVMQEKYEV 1165
            |  ||: | ::     :|| |         :    |||||||||||| |    :  ||:|
Db   1195  FIPVENLETTMRSPVFTDNS---SPPAVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKV 1251

Qy   1166  LVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYLT-GACSRN-YD 1223
           |||||||| | : :||:| || | ||: :| | |: ||::||||||| :|  ||| :  ||
Db   1252  LVLNPSVAATLGFGAYMSKAHGVDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYD 1311

Qy   1224  VIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDEG 1283
            :||||||::||| ::|||| ||   ::  ||||||||||||: : | |||| |: |: |
Db   1312  IIICDECHSTDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSTTG 1371

Qy   1284  TIPPHGKIKEENLKKGRHLIFEATKKHCDELANELARKGITAVSYYRGCDISKIP-EGD 1342
            ||  |||: : ::|||||||| |:|:|||||| :|  :||| ||||| |:| ||   ||
Db   1372  EIPYGKAIPLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGD 1431

Qy   1343  CVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRT 1402
            |||||||| ||:|||||||: |  ::  :|  | ||||||  | : : :|: :|||||||
Db   1432  VVVVATDALMTGFTGDFDSVIDCNTCVIQTVDFSLDPTFTIETTLPQDAVSRTQRRGRT 1491

Qy   1403  GRGRAGIYYVMVDGSCTPSGMVPECNIVEAPDAAKAWYGLSSTEAQTILDTYRTQPGLPAI 1462
           |||:||||: ||: | ||||: |:| ::  :|| |||| : :| |:  | ||  ||||| :
Db   1492  GRGKPGIYRFVAPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLLRAYMNTPGLPVC 1551

Qy   1463  GANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLQLCHQYGYAAAPNDAPRWQG- 1520
              :|:||  :|     :  | :  ::     ||  :|::  ||: |||   | :: :|
Db   1552  QDHLEFWEGVFTGLTHIDAHFLSQTKQSGENFPYLVAYQATVCARAQAPPSWDQMWKCL 1611

Qy   1521  ARLGKK----PCGVLWRLDGADACPGPEPSEV------TRYQM-CFT---EVNTSGTAALA 1567
            |:|     |||: : |  |        |:  :       | |    |       |  ||:
Db   1612  IRLKPTLHGPTPLLYRL------GAVQNEVTLTHPITKYIMTCMSADLEVVTS-TWVLV 1663

Qy   1568  VGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEEI------VEBCASFIPL 1621
            |   :| ||  |  | |: |:| || ::||   | |::: | : :       |:|:  |
Db   1664  GGVLAALAAYCLSTGCVVIVGR------IVLSGKPAIIPDREVLYQEFDEMEBCSQHLPY 1717

Qy   1622  --EAMVAA---IDKLKSTITTTS-------PFTLETALEKLNTFLGPHAATILAIIEYCC 1669
              |:::   |: | |::: :     ||:||  : ::|| | ||  || |:  ::|
Db   1718  IEQGMMLAEQFKQKALGLLQTASRQAEVITP-AVQTNWQKLEVFWAKHMWNFISGIQYLA 1776

Qy   1670  GLVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAG 1729
           || |||:||:|: ::||  |:|:||   | :| | :|| :| :|     || | :||:||
Db   1777  GLSTLPGNPAIASLMAFTAAVTSPLTTGGTLLFNILGGWVAAQLAAPGAATAFVGAGLAG 1836

Qy   1730  TALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVGYSAFNPRAGV 1786
            |:|   : ||   ::| :||||||  :| : |:|:||| |:|:||  |:     | :| |
Db   1837  AAIG---SVGLGKVLVDILAGYGAGVAGLAVAFKIMSGEVPSTEDLVNLLPAILSPGALV 1893

Qy   1787  VGVLSACAMFALTTAGPD---HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEAS 1842
           ||| :::|:  | | ||:    :  ||: :  | :| :| :::: ::  :: : ||:|:|
Db   1894  VGVV--CAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHVVPESDAAARVTAILSSL 1951

Qy   1843  TPWSVISACIRWLHTPTEDDCGLIAWGLEIWQVVCNFPVICFNVLKAGVQSMVNIPGCPF 1902
            |:: :||| |||:||:||| || ::|||: ::  |:|| :|||| : |:|:::|||||:
Db   1952  NWAVRTKLKLTPIAAAGRLDLSGWFTAGYSG---GDIYHSVSHARPRW------FWF 2008

Qy   1903  YSCOKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLC 1962
            ||:||| | | | | |:|||||||| : :::|||:| ||||||| |  | ||| ||| |
Db   2009  VSCQRGYRGVWRGDGIMTRCHCGAEITGHVKNGTMRIV-GPRTCRNMWSGTFPINAYTT 2067

Qy   1963  GSARPDPT-DWTSLVVNVGVRDYCKYEKMGDHIFVTAVSSPNV-CFTQVP--------- 2010
           |  :|||  :|:|:::|  :||: :|::|||| :|::  |:|  || ::
Db   2068  GPCTPLPAPNYKFALWRVSAEEYVEIRRVGDFHYVSGMTTDNLKCPCQIPSPEFFTELDG 2127


Qy   2011  ----------------------GVQVQCYLGPKPTWTTSACCYGPDGKGKTV 2051
                                  ||:|:|||| |||:  ||||||| ||||:||
Db   2128  VRLHRFAPPCKPLLREEVSPVRGLHEYPVVGSQLPC----EPE------PDVAVLTS 2173

Qy   2052  KL--PFRVDGHTPGVRMQLNLRDALETND----------------CNSTNNTPSDEAAVSALVF 2097
            :|   |  |:|||||: :| |||| :||                |:|:|:|||:|||||||:|:
Db   2174  MLTDPSHITAEAAGRRLARGSPPSMASSSASQLSAPSLKATCTANHDSPDAELIEANLLW 2233

Qy   2098  KQEL-------RRTNQLL-----EAISAGVDTTKLPAPSIEEVVRKRQPRARTGSLTLP 2145
            :|:|       |::|::|      ::|:| || | | :: :||:::: |:| |  : ||
Db   2234  RQEMGGNITRVESENKVVILDSFDPLVAEEDEREVSVPA--EILRKSRRFAPALPVWARP 2291

Qy   2146  ------------------------------PPPRSVP----- 2152
                                         |||||||
Db   2292  DYNPLVETWKKPDYEPPVVHGCPLPPPRSPPVPPPRKKRTVVLTESTLSTALAELATKS 2351

Qy   2153  --------GVSCPESLQRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTE 2205
                    |:|    |  || ||||:| ||:|| :  : |:|  | : |  |  |: ::
Db   2352  FGSSSTSGITGDNTTTSSEP--APSGCPPDSDVESYSSMPPLEG--------------- 2393

Qy   2206  TGGGPDDLPSYPPKKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKK 2265
            |||| || |||:   :  ||:|||:|:|||:|:| ||  |::| |:|::       |||:
Db   2394  EPGDPD------------------LSDGSWSTVSS---------GADT------------- 2414

Qy   2266  KLGKSEFSCCSMSYTWTDVISFKTASKVLS-ATRAITSGFLKQRSLVYVTEPRDAELRKQK 2324
            :| :|: ||||||: ||   :|  |::|   :  :| | :|::|:||| |  | |  |||:
Db   2415  ---EDVVCCSMSYSWTGALVTPCAAEEQKLPINALSNSLLRHHNLVYSTTSRSACORQKK 2471

Qy   2325  VTINRQPLFPPSYHKQVRLAKEKASKVVGWMWDYDEVAAHTPSKSAKSHITGLRGTDVRS 2384
            || :: :: ||:|:::|: : :|||| |:|: : ||   :||||:|||  |: |: |||
Db   2472  VTFDRLQVLDSHYQDVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKF-GYGAKDVRC 2530

Qy   2385  GAARKAVLDL----QKCVEAGEIPSHYRQTVIVPKEEVFVKTPQXPKPRLISYPHLE 2440
            :|||||| :     ::||  :  |: ::   ::|:|:|| |:| :|: ||| :||:
Db   2531  -HARKAVAHINSVWKDLLEDSVTPI---DTTIMAKNEVFCVQPEKGGRKPARLIVFPDLG 2586

Qy   2441  MRCVEKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKGRLLSMW--SPDAVGATCDTVCFDS 2497
            :||:|||  : |: :: ||:|: :|||   | |:|  ||| :    |:| |  |||:|||
Db   2587  VRVCEKMALYDVVSKLPLAVMGSSYGFQYSPGQRVEFLVQAWKSKKTPMGPSYDTRCFDS 2646

Qy   2498  TITPBDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYT 2557
            | | :|| :|:|||   ||:||:|| | :||  | ||||: : :|||  |||||:||::|
Db   2647  TVTESDIRTEEAIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGVLT 2706

Qy   2558  TSSSNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVM 2617
            || :|:|||::| :||| :||:::  ::::|||| ||  :||: |::|| |||| :|:::
Db   2707  TSCGNTLTCYIKARAACRAAGLQDCTMLVCGDDLVVICESAGVQEBDAASLRAFTEAMTRY 2766

Qy   2618  GAPQDCVPQPKYSLEELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSA 2677
            ||| | ||:|:| ||::|||||||:  :  ::| :||||||| :||  : | || ||::
Db   2767  SAPPGDPPQPEYDLELITSCSSNVSVAHDGAGKRVYYLTRDPTTPLARAAWETARHTPVN 2826

Qy   2678  AWIGYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNVTVPVEDLPSIIAGV 2737
             ::|:|| : ||||  |:| ||::|:|| |:| || || :: |:|  :  ||||||:  :
Db   2827  SWLGNIIMFAPTLWARMILMTHFFSVLIARDQLEQALNCEIYGACYSIEPLDLPPIIQRL 2886

Qy   2738  HGIEAFSVVRYTNABILRVSQSLLTMTMPPLRAWRKKARAVLASAKRGGAHAKLARFLL 2797
            ||  |||: ||:   |:||: || |:: ||||||::: |:| :|::|| ||: : :|
Db   2887  HGLSAFSLHSYSPGEINRVAACLRKLGVPPLRAWRHRARSVRARLLSRGGRAAICGKYLF 2946

Qy   2798  -WHATSR----PLPDLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVF 2852
                :|    ||||||:|::| || | ::   :  |:  : |: | ||::|:  : ::
Db   2947  NWAVRTKLKLTPIAAAGRLDLSGWFTAGYSG-----GDIYHSVSHARPRW------FWF 2994

Qy   2853  ALGLIAVGLAI 2863
            .||:|| ||.|
Db   2995  CLLLLAAGVGI 3005
```

RESULT 10
US-11-126-662-1
; Sequence 1, Application US/11126662
; Publication No. US20050221297A1

GENERAL INFORMATION:
    APPLICANT: TAKAHAMA, Y.
              SHIRAISHI, J.
    TITLE OF INVENTION: DIAGNOSTIC REAGENT FOR HEPATITIS C VIRUS INFECTION
    NUMBER OF SEQUENCES: 13
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: MORRISON & FOERSTER
        STREET: 2000 Pennsylvania Avenue, NW
        CITY: Washington
        STATE: DC
        COUNTRY: USA
        ZIP: 20006-1888
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Diskette
        COMPUTER: IBM Compatible
        OPERATING SYSTEM: Windows
        SOFTWARE: FastSEQ for Windows Version 2.0b
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/11/126,662
        FILING DATE: 10-May-2005
        CLASSIFICATION: <Unknown>
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: US/08/850,328
        FILING DATE: 02-MAY-1997
        APPLICATION NUMBER: <Unknown>
        FILING DATE: <Unknown>
    ATTORNEY/AGENT INFORMATION:
        NAME: Mays, Thomas D
        REGISTRATION NUMBER: 34,524
        REFERENCE/DOCKET NUMBER: 32273-20004.00
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: 202-887-1500
        TELEFAX: 202-822-0168
        TELEX: 90-4030
    INFORMATION FOR SEQ ID NO: 1:
        SEQUENCE CHARACTERISTICS:
            LENGTH: 3011 amino acids
            TYPE: amino acid
            STRANDEDNESS: single
            TOPOLOGY: linear
        MOLECULE TYPE: peptide
        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-11-126-662-1

Query Match   20.9%; Score 3215.5; DB 20; Length 3011;
Best Local Similarity   29.8%; Pred. No. 4.7e-235;
Matches 964; Conservative 441; Mismatches 1214; Indels 621; Gaps 110;

```
Qy    8 TSPVPAPRTRKNKQTOASYPVSIK----------------TSVERGQRAKRKVQRDARPR   51
          :|:  :|:|||:|            ::|||
Db    3 TNPKP--QKKNKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR   59

Qy   52 NYK----IAGIHDGLQTLAQQAALP------------AHGWGRQDPRHKSRN        86
        |||    ||                              :||                |:|
Db   60 GRRQPIPKARRPEG-RTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRN  118

Qy   87 LGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLF  143
        ||:::|  ||||||:|||  ||||||| | ||       |||||||:|||        :|
Db  119 LGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLPGCSFSIF  177

Qy  144 VVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC-----AD  197
        ::|||  |    ||   ::  :|:  :                   |:|||:|
Db  178 LLALLSCLTVPASAYQVRN-STGLYHVTNDCPNSSIVYEAADAILHTPGCVPCVREGNAS  236

Qy  198 ECWVPANPYI-SHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV--  254
        :||| | || | |||  :     :||||| |||| ||||| : ||:||:|| ||| :
Db  237 RCWVAMTPTVATRDGKLPATQ--LRRHIDLLVGSATLCSALYVGDLCGSVFLVGQLFTFS  294

Qy  255 --RHWLIHIDLNETGTCYLEVPTGIDPGFLGPIGW---MAGKVEAVIFLTKLASQVPYAI  309
          | ||     |  |:| | |||:| |||    ||    :|                :||
Db  295 PRRHW------TTQGCNCSIYPGHITGHR--MAWDMMMWSPTTALVMAQLL-RIPQAI    344

Qy  310 ATMFSSVHYLAVGALIYYASRGKWYQLLALMLY--IEA----------------       346
        | : :  |:  |||:|::|| :|||  || :|:  ||
Db  345 LDMIAGAHWGVLAGIAYFSMVGNWAKVLVVLLFAGVDAETHVTGSAGHTVSGFVSLLA    404

Qy  347 ---------------------------------------TSGNPIRVPTGCSIAE----  362
                                               :|||:     | | | |
Db  405 PGAKQNVQLINTGSWHLNSTALNCNDSLNTGWLAGLFYHHKPNSSGCPERLAS------  458

Qy  363 FCSPL------------------------MIPCPCHSYLSENV-SEVICYSPKWTRPI  395
        :|  |                          |||||  ||:    |:::|  | |
Db  459 -CRPLTDFDQGWGPISYANGSGPDQRPYCWHYPPKPCGIVPAKSVCGPVVCFTPS---PV  514

Qy  396 TLEYNN-----SISWYPYTIPGARGCMVKFKNNT---------WGCCIRRNVPSY-----  436
        |                  | |  |                   |||
Db  515 VVGTTDRSGAPTYSW-----GENDTDVFLNNTRPPLGNWFGCTWMNSTGFTKVCGAPP    568

Qy  437 CTMG----------------------TDAVWNDTRNTYEACCGVTPWLT--------TAWHNGSALKLAILQ  477
        |                         |||::|| :|||  |   ||:|        |||||||  |:| |||
Db  569 CVIGGAGNNTLHCPTDCFRKHPDATYSRCGSGPWLITPRCLVDYPYRLWHYPCTINYTIFK    628

Qy  478 ---YPGSKE--MFKPHNWMSG-HLYFEGSDTPIVVFYDPVNSTLLPPERWARLPGTPPVV  531
           ||| |     | :|     ::  || ||:                   |    |
Db  629 IRMYVGGVEHRLEAACNWTRGERCDLEDRDR-------SELSPLLLTTQWVLPCS----  678

Qy  532 RGSWLQVPQGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVLILLGLC  591
        | |||  |  |  ||    ||||  | |  :|||| |  ||:|::||: :  || |  ||
Db  679 ----FTTLPALSTGLIHLHQNIVDVQYLYGVGSSIASWAIKWEVVLLFLLLA        727

Qy  592 GSKYLILAYLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYL  651
        | |||:    |:| | |
Db  728 DARVCSCLWMMLL----------------------ISQAEAALENLVI-----    753

Qy  652 RCRLRYAALIGFVPMAAGLP--LTFVAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIA  709
        | :||  ||:|:| | | |      || :                  | | |  |: |
Db  754 ---LNAASLAG----THGLVSFLVFFCFA--------------GKWVPGAVY    788

Qy  710 LLVGPWPLVALLTLLHLVTPASAFDTBI---IGGLTIPPVVALVVMSRFGFFAHLLPRCA  766
        |  ||||| |||||| |  ||       |    ||  | :| :  |  | | ||  |  |
Db  789 TFYGMWPL--LLLLLALPQRAYALDTEVAASCGGVVL---VGLMALTLSPYYKRYISWCL  843

Qy  767 LVNSYLWQRWEN-----WFWNVTLRPERFFLVLVCFPGATYDALVTFCVCHVALLC-LTSS  821
        | : ||  | |       | |                |    |:
Db  844 WWLQYFLTRVEAQLHVWIPPLNVRGGRDAVILL---------------        891

Qy  822 AASFFGTDSRVRAHRMLV-------------RLGKCHAWYSHYVLKFFLLVFGEN  863
        | :|    |  | || :|             : |  |  :|    :  | | |||
Db  892 LLAVFGPLWIIQASLLKVPYFVRVQGLLRFCALARMIG-----GHYVQMVIIKLGALT  945

Qy  864 GVFFYKHLH--GDVLPNDFASKLPLQEPFFPEGKARVYRNEGRRLACGDTVDGLPVVAR  921
        |   |: |    | |  |||  :   |||:  :|| :  :  ||| ||| ::|||| ||
Db  946 GTYVYNHLTPLRDWAHNGLRDLAVAVEPVVFSQMETKLITWGADTAACGDIINGLPVSAR  1005

Qy  922 LGDLVFAGLA--MPPDGWAITAPFTLQCLSERGTLSAMAVVMTGIDRTWTGTIFRLGSL  979
        || :: |  |   |  ||: :|  | |: : || |  |:  :::|||  | || ||||
Db 1006 RGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTA  1065

Qy  980 ATSYMGFVCDNVLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDI--YQPPCGAGSLTRC  1037
        | : :|:|| |:  | |||: |||:| |   ||| |     ||    : ||   |||| |
Db 1066 AQTFLATCINGVCWTVYHGAGTRTIASPKGPV--IQMYTNVDQDLVGWPAPQGSRSLTPC  1123

Qy 1038 SCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAA  1097
        :||   |||||| |:: |::| |  :  ||||||:  | |||| ||||| || ||:| ||
Db 1124 TCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAA  1183

Qy 1098 ---RNSGGSVSQIRVRPLVCAGYHPQTAHATLDTKPTVPNEYSVQLLIAPTGSGKSTKL  1154
           :|:   :  ||   ||||||| || |||  |:  |    : | : ||||||||||:
Db 1184 VCTRGVAKAVDFIPVENLETTMRSPVFTDNS--SPPVVPQSFQVAHLHAPTGSGKSTKV  1240

Qy 1155 PLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGYL  1214
        |  :  :|: ||||||||| |  : :|:   ||::|    :|  || ||: |||||| :|
Db 1241 PAAYAAQGYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGSPITYSTYGKFL  1300

Qy 1215 T-GACSRN-YDVIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHA  1272
```

```
Db  1301  ADGGCSGGAYDIIICDECHSTDRATSILGIGTVLDQAETAGARILVVLATATPPGSVTVHP  1360
          ||      |||||||||||||    |||     ||          ||||||||
Qy  1273  NITEIQLTDEGTIPFHGKCIKEENLKKGRHLIFEATKKHCDELANELARKGITAVSYRG    1332
          ||:  |||| |  |||| |  |  ||||| ||| |  ||||||||| |    ||| |||
Db  1361  NIEEVALSTTGEIPFYGKAIPLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRG   1420
Qy  1333  CDISKIP-EGDCVVATDALCTGYTGPDFPDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVS  1391
          || | |   ||||||||||  ||||| | |||  |    ||  ||||||||| | |
Db  1421  LDVSVIPTSGDVVVVATDALMTGYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETITLPQD   1480
Qy  1392  AIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTILD   1451
          ||: |||||||||||| |||| ||| ||||  || |||||    | | |   |  ||
Db  1481  AVSRTQRRGRTGRGKPGIYRFVAPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLR   1540
Qy  1452  TYRTQPGLPAIGANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLQLCHQYGYA   1510
          : | ||||| |  |::: ||  :  :::   : | :|  ||  ||    || | |   |
Db  1541  AYMNTPGLPVCQDHLEEWEGVFTGLTHIDAHFLSQTKQSGENLPYLVAYQATVCARAQAP   1600
Qy  1511  APNDAPRWQG-ARLGKK---PCGVLWRLDGADACPGPEPSE-----VTRYQM-CFT---  1556
          ||      ||  ||  ||    | | |||  |  ||  ||          : | :  |||
Db  1601  PPSWDQMWKCLIRLKPTLHGPTPLLYRL--------GAVQNEITLTHPVTKYIMTCMSADL 1653
Qy  1557  EVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSIISVPTGATVAPVVDEEEI-----  1611
          ||:|| |||||| | | |: :|| || |    ||| || |  | | | |||
Db  1654  EVVTS-TWVLVGGVLAALAAYCLSTGCVVIVGR------VVLSGKPAIIPDREVLYRBFD   1706
Qy  1612  -VEECASFIPL--EAMVAA--IDKLKSTITTS-------PPTLETALEKLNTFLGPHA    1658
          :|||| | |||  |  ||  ||    |  |  |        |||  | ||| | | |
Db  1707  EMEECSQHLPYIEQGMMLAEQFKQKALGLLQTASRQABVIAP-AVQTNWQKLETFWAKHM  1765
Qy  1659  ATILAIIEYCCGLVTLPDNPFASCVFAPIAGITTPLPHKIKMFLSLFGGAIASKLTDARG  1718
          |  | |  || ||:||| || |:|: ||| || || | |: :|| || |:|  | |
Db  1766  WNFISGIQYLAGLSTLPGNPAIASLMAFTAAVTSPLTTSQTLLFNILGGWVAAQLAAPGA  1825
Qy  1719  ALAFMMAGAAGTALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGL  1775
          | |  : | | | |      |     ||:|||||||     |  |  | | | | ||::|
Db  1826  ATAFVGVGALGLAGAIG---SVGLGKVLIDILAGYGAGVAGALVAFKIMSGEVPSTEDLVNL 1882
Qy  1776  VYSAFNPAAGVGVLSACAMFALITTAGPD----HWPNRLLTMLARSNTVCNEYFIATRDI  1831
          : || || | :|| ||  | || |  |::     ||  |  | :    : | |  |  |
Db  1883  LPAILSPGALVVGVV--CAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDA   1940
Qy  1832  RRKILGILEASTPWSVISACIRWLHTPTEDDCGLIAWGLEIWQYVCNFFVICFNVLRAGV  1891
          |   |     | | | |||| ||||| |      | :||:| |      : | |  :|
Db  1941  AARVTAILSSLIVTQLLRRLHQWISSECTTPCS-GSWLRDIWDWICEVLSDFKTWLKA--  1997
Qy  1892  QSMVNIPGCGPFYSCQKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYW    1951
          |   | |    ||||:||  | || | ||| || |  |   :|  ||||||| | |:|
Db  1998  KLMPQLPGIPFVSCQRGYKGVWRGDGIMHTRCHCGAEITGHVKNGTMRIV-GPRTCRNMW   2056
Qy  1952  RGAVPVNARLCGSARPDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFTQV  2009
          |||      || |:| |   | ||   ||| || ||  |||  | ||  |   |  | | |
Db  2057  SGTFPINAYTTGPCTPLPAPNYTFALWRVSAEEYVEIRQVGDFHYVTGMTTDNLKCPCQV  2116
Qy  2010  P---------------------PTLRAAVAVD-------GVQVQCYLGEPKTPWTTSAC   2040
          |                     || ||  |         ||  |  | ||  |   :|
Db  2117  PSPEFFTELDGVRLHRFAPPCKPLLREEVSFRVGLHEYPVGSQLPC---EPE-----    2165
Qy  2041  CYGPGDKGKTVKL--PFRVDGHTPGVRM---------QLNLRDALETNDCNSTNN    2084
          || |||| ||||   |||| |  || ||          |||| | |  ||  | |
Db  2166  ---PDVAVLTSMLTDPSHITAEAAGRRLARGSPPSVASSASQLSAPSLKAT--CTANHD   2220
Qy  2085  TPSDEAAVSALVFKQEL------RRTNQLL-----EAISAGVDTTKLPAPSIEEVVRK   2132
          :||| |  ||    ||       || |:|      ||||| ||| |||| |:||   |
Db  2221  SPPAELIEANLLWRQEMGGNITRVESENKVVILDSFDPLVAEEDEREISVPA--EILRKS  2278
Qy  2133  RQFRARTGSLTLPPPPRSVPGVSCP--ESLQRSD--------PLEGPSNLPFSPP----  2177
          |     |||| |||||||||| ||   |||| |         || |||||| ||
Db  2279  RRF------AQALPVWAR--PDYNPLVETWKKPDYEPPVVHGCPLPPKSPPVPPPRKKR   2331
Qy  2178  --VL--QLAMEMPLLGAGECNPFTAIGCAMETGGGPDDLPS-YPPKKEVSEW-------  2225
          : |    || |||||||| ||||| || |||||||||||| |:| ||||| |
Db  2332  TVLTESTLSTALAELATRSFGSSTSGITGDNTTTSSEPAPSGCPPDSAESYSSMPPL    2391
Qy  2226  --------SDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKKLGKSEFSCSM   2276
          :        ||| |    |  | ||||| |||| ||  | |  ||     ||
Db  2392  EGEPGDPDLSDGSWNSTVSSEAN---------AEDVVCCSM                    2422
Qy  2277  SYTWTDVISFKTASKVLS-ATRAITSGFLKQRSLVYVTEPRDAELRKQKVTINRQPLFPP  2335
          || |  :: |  |  | |  ||| : ||: |:|||| |  | |  |:|||| : | |  |
Db  2423  SYSWTGALVTPCAEEQKLPINALSNSLLRHHNLVYSTTSRSACQRQKKVTFDRLQVLDS   2482
Qy  2336  SYHKQVRLAKEKASKVVGVWMDYDEVAAHTPSKSAKSHITGLRGTDVRSGAARKAVLDL-  2394
          ::: ||  ||:||| ||  ::: ||  || |   |  |   |   ||| | |||| | :
Db  2483  HYQDVLKEVKAAASKVKANLLSVEEACSLITPPHSAKSKF-GYGAKDVRC-HARKAVTHIN 2540
Qy  2395  ---QKCVEAGEIPSHYRQTVIVPKEEVVKTPQKTKKPTKKPPRLISYPHLEMRCVEKMYYGQ 2451
             :| ||  |  |:  |  :: | :: |    |  || |   ||| | |  | ||||||| |
Db  2541  SWKDLLEDNVTPI---DTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYD   2597
Qy  2452  VAPDVVKAVMGDAYGF-VDPRTRVKRLLSMW--SPDAVGATCDTVCFDSTITPEDIMVET  2508
          | |:||||:|| : || |||||  ||||| |   |||| |||| |||| |||:|||: ||
Db  2598  VVTKLPLAVMGSSYGFQYSPGQRVEFLVQAWKSKKTPMGFSYDTRCFDSVTESDIRTEE   2657
Qy  2509  DIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTTSSSNSLTCWL  2568
          ||:  | | |: ||  | | |:||||||   |||   |||||: |||: ||   | ||:|
Db  2658  AIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRAASGVLTTSCGNTLTCYI  2717
Qy  2569  KVNAAAEAQGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMGAPQDCVPQPK  2628
          | :|| | |  |    |:|||| ||:| ||| :|: || |||  | | | || ||:| |
Db  2718  KARAACRAAGLQDCTMLVCGDDLVVICESAGVQEDAASLRAFTEAMTRYSAPPGDPPQPE  2777
Qy  2629  YSLEELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYP  2688
          | ||:|||||||| | | ||| | | |||||  ||  : ||  | |  |||:| |||
Db  2778  YDLELITSCSSNVSVAHDGAGKRVYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMFAP  2837
Qy  2689  CLWVSRVLAVHFMEQMLFEDKLPETVTFDMYGKNYTVPVEDLPSIIAGVHGIEAFSVVRY  2748
           ||   ::  ||   :| :  |  :: |: || ||::  :||| ||  |||:||||:: |
Db  2838  TLWARMILMTHFFSVLIARDQLEQALDCEIYGACYSIEBPLDLPPIIQRLHGLSAFSLHSY  2897
Qy  2749  TNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-WHATSR----  2803
          : | |:||: || | || ||||||: || | | ::|||| || | |:| || | :
Db  2898  SPGEINRVAACLRKLGVPPLRAWRHRARSVRARILLARGGRAAICGKYLFNWAVRTKLKLT  2957
Qy  2804  PLPDLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVPALGLIAVGLAI  2863
           | :    :   |  : : :  |  |  :       || |:|   ::  | :||| |:|
Db  2958  PIRAAAGQLDLSGWFTAGYSG-----GDIYHSVSHARPRW---IWFCLLLLAAGVGI    3005
```

RESULT 11
US-09-952-572-9
; Sequence 9, Application US/09952572
; Patent No. US20020119495A1
; GENERAL INFORMATION:
;   APPLICANT: HAWAII BIOTECHNOLOGY GROUP, Inc.
;   APPLICANT: NAKANO, Eileen
;   APPLICANT: CLEMENTS, David
;   APPLICANT: HUMPHREYS, Tom
;   TITLE OF INVENTION: IMMUNOGENIC COMPOSITION OF HEPATITIS C
;   FILE REFERENCE: HAWBIO1100
;   CURRENT APPLICATION NUMBER: US/09/952,572
;   CURRENT FILING DATE: 2001-09-13
;   PRIOR APPLICATION NUMBER: US 60/230,927
;   PRIOR FILING DATE: 2000-09-13
;   NUMBER OF SEQ ID NOS: 9
;   SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
;   LENGTH: 3011
;   TYPE: PRT
;   ORGANISM: Hepatitic C Virus
US-09-952-572-9

Query Match   20.9%;   Score 3214.5;   DB 9;   Length 3011;
Best Local Similarity   29.4%;   Pred. No. 5.6e-235;
Matches 956;   Conservative 440;   Mismatches 1213;   Indels 641;   Gaps 107;

```
Qy   875  VLPNDFASKLPLQEP--FPPFEGKARVYRNEGRRLACGDTVDGLPVVARLGDLVFAGLA-   931
Db   959  WAHNGLRDLAVAVEPVFSRMETKLITWGAD--TAACGDIINGLPVSARRGQEILLGPAD   1016

Qy   932  -MPPDGWAITAPFTLQCLSRGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSYMGFVCDN   990
Db  1017  GMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTATQTFLATCING  1076

Qy   991  VLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDI--YQPPCGAGSLTRCSCGETKGYLVT  1048
Db  1077  VCWTVYHGAGTRTIASPKGPV--IQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVT  1134

Qy  1049  RLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAA---RNSGGSVS  1105
Db  1135  RHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAVCTRGVAKAVD  1194

Qy  1106  QIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQLLIAPTGSGKSTKLPLSYMQEKYEV  1165
Db  1195  FIPVENLGTTMRSPVFTDNS---SPPAVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKV  1251

Qy  1166  LVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSYGMYLT-GACSRN-YD  1223
Db  1252  LVLNPSVAATLGFGAYMSKAHGVDNIRTGVRTIITGSPITYSTYGKFLADGGCSGGAYD  1311

Qy  1224  VIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDEG  1283
Db  1312  IIICDECHSTDATSIGIGTVLDQAETAGARLVVLATATPPGSVTVSHPNIEEVALSTTG  1371

Qy  1284  TIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELARKGITAVSYRGCDISKIP-EGD  1342
Db  1372  EIPFYGKAIPLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGD  1431

Qy  1343  CVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRT  1402
Db  1432  VVVVSTDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTLPQDAVSRTQRRGRT  1491

Qy  1403  GRGRAGIYYYDGSCTPSGMVPECNIVEAFDAAKAMWYGLSSTEAQTILDTYRTQPGLPAI  1462
Db  1492  GRGKPGIYRFVAPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVC  1551

Qy  1463  GANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLCHQYGYAAPNDAPRWQG-  1520
Db  1552  QDHLEFWEGVFTGLITHIDAHFLSQTKQSGENFPYLVAYQATVCARAQAPPSWDQMWKCL  1611

Qy  1521  ARLGKK--PCGVLWRLDGADACPGPEPSEV------TRYQM-CFT--EVNTSGTAALA  1567
Db  1612  IRLKPTLHGPTPLLYRL------GAVQNEVTLTHPITKYIMTCMSADLEVVTS-TWVLV  1663

Qy  1568  VGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI-----VEECASFIPL  1621
Db  1664  GGVLAALAAYCLSTGCVVIVGR------IVLSGKPAIIPDREVLYQEFDEMEECSQHLPY  1717

Qy  1622  --EAMVAA---IDKLKSTITTTSPF----------TLETALEKLNTFLPGPHAATILAIIEYCCG  1670
Db  1718  IEQGMMLAEQFKQKALGLLQTASRHAEVITPAVQTNWQKLEVFWAKHMWNFISGIQYLAG  1777

Qy  1671  LVTLPDNPFASCVFAFIAGITTLPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAGT  1730
Db  1778  LSTLPGNPAIASLMAFTAAVTSPLTTGQTLLFNILGGWVAQLAAPGAATAFVGAGLAGA  1837

Qy  1731  ALGTWTSVGF---VFDMLGGVAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVV  1787
Db  1838  AIG---SVGLGKVLVDILAGYGAGVAGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVV  1894

Qy  1788  GVLSACAMFALTTAGPD----HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEAST  1843
Db  1895  GVV--CAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAARVTAILSSLT  1952

Qy  1844  PWSVISACIRWLHTPTEDDCGLIAWGLEIWQYVICFNVLKAGVQSMNIPGCGPFY  1903
Db  1953  VTQLLRRLHQWISSECTTPCS-GSWLRDIWDWICEVLSDFKTWLKA--KLMPQLPGIPFV  2009

Qy  1904  SCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCG  1963
```

```
Qy     8  TSPVPAPRTRKNKQTQASYPVSIK----------------TSVERGQRAKRKVQRDARPR    51
Db     3  TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR   59

Qy    52  NYK----IAGIHDGLQTLAQAALP-----------------AHGWGRQDPRHKSRN     86
Db    60  GRQPIPKARRPEG-RTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRN   118

Qy    87  LGILLDYPLGWIGDVTHTPLVGPIVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLF   143
Db   119  LGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLPGCSFSIF  177

Qy   144  VVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC-----AD  197
Db   178  LLALLSCLTVPASAYQVRN-SSGLYHVTNDCPNSSIVYEAADAILHTPGGVPCVREGNAS  236

Qy   198  ECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV--  254
Db   237  RCWVAVTPTVATRDGKLPFTQ-LRRHIDLLVGSATLCSALYVGDLCGSVPLVGQLFTFSP  295

Qy   255  -RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGW---MAGKVEAVIFLTKLASQVPYAIA  310
Db   296  RRHW-------TTQDCNCSIYPGHITGHR--MAWDMMMNWSPTAALVVAQLL-RIPQAIM  345

Qy   311  TMFSSVHYLAVGALIYYASRGKWTYQLLALMLY------IEATSGNPIRV---------  354
Db   346  DMIAGAHWGVLAGIAYFSMVGNWAKVLVVLLLFAGVDAETHVTGGNAGRTTAGLVGLLTP  405

Qy   355  --------------------------------------PTGC--------------SI  360
Db   406  GAKQNIQLINTNGSWHINSTALNCNESLNTGWLAGLFYQHKFNSSGCPERLASCRRLTDF  465

Qy   361  AEFCSPLMI---------------PCPCHSYLSENV-SEVICYSPKWTRPITLEYNN--  401
Db   466  AQGWGPISYANGSGLDERPYCWHYPPRPCGIVPAKSVCGPVYCFFPS---PVVVGTTDRS  522

Qy   402  ---SISWYPYTIPGARGCMVKFKNNT------WGCCRIRNVPSY-------CTMG----  440
Db   523  GAPTYSW------GANDTDVFVLNNTRPPIGNWFGCTWMNSTGFTKVCGAPPCVIGGVGN  576

Qy   441  -------TDAVWNDTRNTYEACGVTPWLT--------------TAWHNGSALKLAILQ--  482
Db   577  NTLLCPTDCFRKHPEATYSRCGSGPWITPRCMVDYPYRLMHYPCTINYTIFKVRMYVGGV  636

Qy   483  E--MPKPHNWMSG--HLYFEGSDTPIVFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVP  539
Db   637  EHRLEAACNWTRGERCDLEDRDR-----SELSPLLLSTTQWQVLPCS-----  678

Qy   540  QGFYSDVKDLATGLITKDRAWKNYQVLYSATGALSLTGVTTKAVVLILGLCGSKYLILA  599
Db   679  ---FTTLPALSTGLIHLHQNIVDVDVQYLYGVGSSIASWAIKWEYVVLLFLLLADARVCSCL  735

Qy   600  YLCYLSLCFGRASGYPLRRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAA  659
Db   736  WMMLL--------------------ISQAEAALENLVI-------LNAAS  758

Qy   660  LLGFVPMAAGLP--LTFFVAAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALLVGPWPL  717
Db   759  LAG----THGLVSFLVFFCFA--------WYLK-----GRWVPGAVYALYGMWPL  796

Qy   718  VALLTLLHLVTPASAFDTEI---IGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQ  774
Db   797  --LLLLLALPQRAYALDTEVAASCCGGVVL----VGLMALTLSPYYKYRYISWCMWWLQYFLT  851

Qy   775  RWEN---WFWNVTLRPERRFLVLVCFPGATYDALVTFCVCHVALLC-LTSSAASFFGT-  828
Db   852  RVEAQLHVWVPPLNVRGGRDAVILL------------MCVVHPTLVFDITKLLLAIFGPL  899

Qy   829  --------------DSRVRAHRMLVRLGKCHAWYSHYVLKFLLVFGENGVFFYKHLH--GD  874
Db   900  WILQASLLKVPYFVRVQGLLRICALARKIAG-GHYVQMAIIKLGALTGTYVYNHLTPLRD  958
```

Db   2996 LLLLAAGVGI 3005

RESULT 12
US-09-747-419-20
; Sequence 20, Application US/09747419
; Patent No. US2002015582A1
; GENERAL INFORMATION:
; APPLICANT: Lemon, Stanley
; APPLICANT: Yi, Minkyung
; TITLE OF INVENTION: REPLICATION COMPETENT HEPATITIS C VIRUS AND METHODS OF USE
; FILE REFERENCE: 265.0007 0101
; CURRENT APPLICATION NUMBER: US/09/747,419
; CURRENT FILING DATE: 2000-12-23
; PRIOR APPLICATION NUMBER: US 60/171,909
; PRIOR FILING DATE: 1999-12-23
; SOFTWARE: PatentIn version 3.0
; NUMBER OF SEQ ID NOS: 34
; SEQ ID NO 20
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polyprotein
US-09-747-419-20

Query Match   20.9%; Score 3214.5; DB 9; Length 3011;
Best Local Similarity   29.4%; Pred. No. 5.6e-235;
Matches 956; Conservative 440; Mismatches 1213; Indels 641; Gaps 107;

Qy   8 TSPVPAPRTRKNKQTQASYPVSIK---------------------TSVERGQRAKRKVQRDARPR   51
           ::|: :|:: |:::   ::|
Db   3 TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR   59

Qy   52 NYK----IAGIHDGLQTLAQAALP-----------------------AHGWGRQDPRHKSRN   86
              :   ::|:
Db   60 GRRQPIPKARRPEG-RTWAQPGYPWPLYGNEGCGWAGWLLSPRGGSRPSWGPTDPRRRSRN 118

Qy   87 LGILLDYPLGWIGDVTTHTPLVGPLVGAVVRPVCQIVRLLEDGVNWATG--WFGVHLF 143
           |:::| ::|
Db   119 LGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLPGGCSFSIF 177

Qy   144 VVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGGVIC-----AD 197
                  : ||| ||
Db   178 LLALLSCLTVPASAYQVRN-SSGLYHVTNDCPNSSIVYEAADAILHTPGGCVPCVREGNAS 236

Qy   198 ECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV--- 254
           || :
Db   237 RCWVAVTPTVATRDGKLPTTQ-LRRHIDLLVGSATLCSALYVGDLLCGSVFLVGQLFTFSP 295

Qy   255 -RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGW--MAGKVEAVIFLTKLASQVPYAIA 310
            :||| :::::
Db   296 RRHW-----TTQDCNCSIYPGHITGHR--MAWDMMWSPTAALVVAQLL-RIPQAIM 345

Qy   311 TMFSSVHYLAVGALIYYASRGKWYQLLLALMLY------IEATSGNPIRV--------- 354
           :::   :::
Db   346 DMIAGAHWGVLAGIAYFSMVGNWAKVLVVLLLFAGVDAETHVTGNAGRTTAGLVGLLTP 405

Qy   355 --PTGC-------SI 360
                                                                  ||
Db   406 GAKQNIQLINTNGSWHINSTALNCNESLNTGWLAGLFYQHKFNSSGCPERLASCRRLTDF 465

Qy   361 AEFCSPLMI----------------------------PCPCHSYLSENV-SEVICYSPKWTRPITLEYNN-- 401
           | ||:| :::
Db   466 AQGWGPISYANGSGLDERPVCWHYPPRPCGIVPAKSVCGPVYCFTPS---PVVVGTTDRS 522

Qy   402 --SISWYPYTIPGARGCMVKFKNNT------WGCCRIRNVPSY------CTMG----- 440
             : || :::
Db   523 GAPTYSW-------GANDTDVFVLNNTRPPLGNWFGCTWMNSTGFTKVCGAPPCVIGGVGN 576

Qy   441 ------TDAVWNDTRNTYEACGVTPWLT------TAWHNGSALKLAILQ---YPGSK 482
                 :   :
Db   577 NTLLCPTDCFRKHPEATYSRCGSGPWITPRCMVDYPYRLWHYPCTINYTIFKVRMYVGGV 636

Qy   2010 SCQRGYRGVWRGDGIMHTRCHCGAEITGHVKNGTMRIV-GPRTCRNMWSGTFPINAYTTG 2068
           |||:|| | | ||||| |::|
Db   1964 SARPDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFTQVP----------- 2010

Qy   2069 PCTPLPAPNYKFALWRVSAEEVYEIRRVGDFHVSGMTDNLKCPCQIPSPEFFTELDGV 2128
           :|:::   ::
Db   2011 ------PTLRAAVAVD-------GVQVQCYLGEPKTPWTTSACCYGPDGKGTVK 2052

Qy   2129 RLHRFAPPCKPLLREEVSFRVGLHEYFVGSQLPC---EPE---------PDVAVLTSM 2174
           |::|| ::|
Db   2053 L--PFRVDGHTPGVRMQLNLRDALETND---------CNSTNNTPSDEAAVSALVFK 2098

Qy   2175 LTDPSHITAEAAGRRLARGSPPSMASSSASQLSAPSLKATCTANHDSPDAELIEANLLWR 2234
                                                         ::
Db   2099 QEL-------RRTNQLL-----EAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLP- 2145

Qy   2235 QEMGGNITRVESENKVVILDSFDPLVAEEDEREVSVPA--EILRKSRFRARALPVWARPD 2292
           |:::   :::
Db   2146 PPPRSVP--- 2152

Qy   2293 YNPPLVETWKKPDYEPPVVHGCPLPPPRSPPVPPRRKKRTVVLTESTLSTALAELATKSF 2352
                                                   :::
Db   2153 -----GVSCPESLQRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTET 2206

Qy   2353 GSSSTSGITGDNTTTSSEP--APSGCPPDSDVESYSSMPPLEG---------------E 2394
           |::| ||
Db   2207 GGGPDDLPSYPPKKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKK 2266

Qy   2395 -LSDGSWSTVSS---------------------------GADT------------ 2414
            : ||
Db   2267 LGKSEFSCCSMSYTWTDVISFKTASKVLS-ATRAITSGFLKQRSLIVYVTEPRDAELRKQKV 2325

Qy   2415 --EDVCCSMSYSWTGALVTPCAAEEQKLPINALSNSLLRHHNLVYSTTSRSACQRQKKV 2472
             :|| ||
Db   2326 TINRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSG 2385

Qy   2473 TFDRLQVLDSHYQDVLKEVKAAASKVKANLLSVEEACSLTPHSAKSKF-GYGAKDVRC- 2530

Qy   2531 ---DTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGV 2587
Db   2386 AARKAVLDL---QKCVEAGEIPSHYRQTVIVPKEEVVKTPQKPTKKPPRLISYPHLEM 2441

Db   2473 TFDRLQVLDSHYQDVLKEVKAAASKVKANLLSVEEACSLTPHSAKSKF-GYGAKDVRC- 2530

Qy   2588 RCVEKMYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMW---SPDAVGATCDTVCFDST 2498
Db   2531 HARKAVAHINSVWKDLLEDSVTPI---DTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGV 2587

Qy   2499 ITPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGMIAYDGREIGYRRCRSSGVYTT 2558
Db   2588 RVCEKMALYDVVSKLPLAVMGGSSYGFQYSPGQRVEFLVQAWKSKKTPMGFSYDTRCFDST 2647

Qy   2559 SSSNSLTCWLKVNAAAEQAGMKNPRFLICGGDCTVIWKSAGADADKQAMRVFASWMKVMG 2618
Db   2648 VTBSDIRTEEAIYQCCDLDPQARVAIKSLTERLLYVGGPLTNSRGENCGYRRCRASGVLTT 2707

Qy   2619 APQDCVPQPKYSLEELTSCCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAA 2678
Db   2708 SCGNTLTCYIKARAACRAAGLQDCTMLVCGDLVVICESAGVQEDAASLRAFTEAMTRYS 2767

Qy   2679 WIGYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKGNYTVPVEDLPSIIAGVH 2738
Db   2768 APPGDPPQPEYDLELITSCCSSNVSVAHDGAGKRVYYLTRDPTTPLARAAWETARHTPVNS 2827

Qy   2739 GIEAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARPLL- 2797
Db   2828 WLGNIIMFAPTLWARMLMTHFSVLIARDQLEQALNCEIYGACYSIEPLDLPPIIQRLH 2887

Qy   2798 WHATSR-----PLPDLDKTSVARYTTFNVCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFA 2853
Db   2888 GLSAFSLHSYSPGEINRVAACLRKLGVPPLRAWRHRARSVRARLLSRGGRAAICGKYLFN 2947

Qy   2854 LGLIAVGLAI 2863
Db   2948 WAVRTKLKLTPIAAAGRLDLSGWFTAGYSG-----GDIYHSVSHARPRW-------FWFC 2995

```
Qy   483  E--MFKPHNWMSG-HLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVP   539
                :    :||                                  |||
Db   637  EHRLEAACNWTRGERCDLEDRDR-------SELSPLLLSTTQWQVLPCS------        678

Qy   540  QGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGLCGSKYLILA   599
          :||  :::|   |: ||         ||    ::|   |  ::| ||||:||||||||:|
Db   679  ---FTTLPALSTGLIHLHQNIVDVQYLYGVGSSIASWAIKWEYVVLLPLLLADARVCSCL   735

Qy   600  YLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAA   659
           :                                :|:| :||
Db   736  WMMLL------------ISQAEAALENLVI------LNAAS                     758

Qy   660  LLGFVPMAAGLP--LTFFVAAAAAQPDYDWWVRLLVAGLVLWAGRNRGHRIALLVGPWPL   717
          | |   :   | :   |:                    |||:|             ||||
Db   759  LAG----THGLVSFLVFFCFA-------WYLK--------GRWVPGAVYALYGMWPL     796

Qy   718  VALLTLLHLVTPASAFDTEI---IGGLTIPPVALVVMSRFGFFAHLLPRCALVNSYLWQ    774
           ||| |::   ||:|                 |||||:|:                |:|
Db   797  --LLLLALPQRAYALDTEVAASCGGVVL---VGLMALTLSPYYKRYISWCMWWLQYPLT    851

Qy   775  RWEN----WFWNVTLRPERFFLVLVCFPGATYDALVTFCVCHVALLC-LTSSAASFFGT-  828
                   ||:|  |  |::  :||                    :|
Db   852  RVEAQLHVWNVPPLNVRGGRDAVILL--------MCVVHPTLVFDITKLLLLAIFGPL    899

Qy   829  -------DSRVRAHRMLVRLGKCHAWYSHYVLKFFLLVFGENGVFFYKHLH--GD      874
               :||:| ||:  |||:| |:   |  |||  ::    |  :::|:  ||
Db   900  WILQASLLKVPYFVRVQGLLRICALARKIAG-GHVVQMAIIKLGALTGTYVYNHLTPLRD   958

Qy   875  VLPNDFASKLPLQEP--FPPFEGKARVYRNEGRRLACGDTVDGLPVVARLGDLVFAGLA    931
           | |||     | |:   |||||||||||   ||||||||:|||||| |||||:|||||
Db   959  WAHNGLRDLAVAVEPVVFSRMETKLITWGAD--TAACGDIINGLPVSARRGQEILLGPAD  1016

Qy   932  -MPPDGWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSYMGFVCDN  990
              ||||   ||       ||||  :|   ||||||||:||||||||||||  ||||:
Db  1017  GMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQISTATQTFLATCING   1076

Qy   991  VLYTAHHGSKGRRLAHPTCSIHPITVDAANDQDI--YQPPCGAGSLTRCSCGETKGYLVT  1048
          | ||::||  |||| :|| :|:| |:  | ||::   | | |:|||| | |||  ||||||
Db  1077  VCWTVYHGAGTRTTIASPKGPV--IQMYTNVDQDLVGWPAFQGSRSLTPCTCGSSDLYLVT  1134

Qy  1049  RLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAA---RNSGGSVS  1105
          | |::   :   |  |||||||||||||||||:|: | |||||:||||||   || ||| 
Db  1135  RHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAVCTRGVAKAVD   1194

Qy  1106  QIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQLLIAPTGSGKSTKLPLSYMQEKYEV  1165
          :| ||  | ||||   || ||  |||:|  : ||:||||||||||||||||: ||||:|:
Db  1195  FIPVENLGTTMRSPVFTDNS---SPPAVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKV  1251

Qy  1166  LVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGASLTYSTYGMYLT-GACSRN-YD  1223
          |||||||| | : || | |||| ||| |  |:  : | :::|||||| :|  | | :  ||
Db  1252  LVLNPSVAATLGFGAYMSKAHGVDPNIRTGVRTIITGSPITYSTYGKFLADGGCSGGAYD  1311

Qy  1224  VIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDEG  1283
          :|||||||:||||::||||:|| :|   ||||||||||||||:  | |||| |: ::|
Db  1312  IIICDECHSTDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVSHPNIEEVALSTTG  1371

Qy  1284  TIPFHGKKIKEENLKKGRHLIFEATKKHCDELANELARKGITAVSYYRGCDISKIP-EGD  1342
          ||||:||  | ::||||||||| |:|| |||| |  |  ||| | ||||||:| :| :||
Db  1372  EIPFYGKAIPLEVIKGGRHLIFCHSKKKCDELAAAKLVALGINAVAYYRGLDVSVIPTSGD  1431

Qy  1343  CVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFMGVRVCGVSAIVKGQRRGRT   1402
          ||||::||| |||:|||||| ||   |  |||||:|||| | | : | ||:|||||||||
Db  1432  VVVVSTDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTLPQDAVSRTQRRGRT  1491

Qy  1403  GRGRAGIYYYVDGSCTPSGMVPECNIVEAPDAAKAWYGLSSTEAQTILDYRTQPGLPAI   1462
          |||:  |||: | | :|| || ||| :|||:   ||| | ::|  |  |   :||||| 
Db  1492  GRGKPGIYRFVAPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVC  1551

Qy  1463  GANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLQLCHQYGYAAPNDAPRWQG-  1520
          ::: :||     :|   :  || : :| |::| |  | |  ||::|||||| :| | |
Db  1552  QDHLEFWEGVFTGLTHIDAHFLSQTKQSGENFPYLVAYQATVCARAQAPPSWDQMWKCL   1611
```

```
Qy  1521  ARLGKK---PCGVLWRLDGDADACPCGPEPSEV------TRYQM-CFT---EVNTSGTAALA  1567
           | | |:    |::| :                     || :| |::     |:|  ||:
Db  1612  IRLKPTLHGPTPLLYRL------GAVQNEVTLTHPITKYIMTCMSADLEVVTS-TWVLV    1663

Qy  1568  VGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI------VEECASFIPL    1621
          :|| ||: | ||      ||  |    :|| |||    |:::         |||    ||
Db  1664  GGVLAALAAYCLSTGCVVIVGR------IVLSGKPAIIPDREVLYQEFDEMEECSQHLPY   1717

Qy  1622  --EAMVAA---IDKLKSTITTSPF------TLETALEKLNTFLGPHAATLAIIEYCCG    1670
            | :  :    |:|          | : |  | | :   : ||   :|| |  : | |
Db  1718  IEQGMMLAEQFKQXALGLLQTASRHAEVITPAVQTNWQKLEVFWAKHMWNFISGIQYLAG  1777

Qy  1671  LVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAGT   1730
          | |||:||  || : | ||:||||   :: | ::|||:::::|| |  ||||: |||
Db  1778  LSTLPGNPAIASLMAFTAAVTSPLTTGQTLLFNILGGWVAAQLAAPGAATAFVGAGLAGA  1837

Qy  1731  ALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVVSAFNPAAGVV  1787
          |:|        |  : |:|||||| | | : |  :  ||| : :|| |||||    ||||
Db  1838  AIG---SVGLGKVLVDILAGVGAGVAGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVV  1894

Qy  1788  GVLSACAMFALTTAGPD----HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILLEAST  1843
          ||: :| ||||:|||      ||:|||| ::|||:  |::   ::||:|| :| |::|::|
Db  1895  GVV--CAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAARVTAILSSLT  1952

Qy  1844  PWSVISACIRWLHTPTEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFY   1903
          ||:::||  ||| ||      | :  :|:| | |::  :|  ||  |  : :| |||||
Db  1953  VTQLLRRLHQWISSECTTPCS-GSWLRDIWDWICEVLSDFKTWLKA--KLMPQLPGIPFV   2009

Qy  1904  SCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCG  1963
          ||| || | | | ||: |||| |||     ::|  || |||||| | |  || || :|||
Db  2010  SCQRGYRGVWRGDGIMHTRCHCGAETGHVXNGTMRIV-GPRTCRNMWSGTFPINAYTTG   2068

Qy  1964  SARPDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFTQVP-------     2010
           :| |   :: :  |:   :                    :      : |
Db  2069  PCTPLPAPNYKFALMRVSAEEYVEIRRVGDFHYVSGMTTDNLKCCPCQIPSPEFFTELDGV  2128

Qy  2011  ------------PTLRAAVAVD---------GVQVQCYLGEPKTPWTTSACCYGPDGKGKTVK  2052
                     |  :  | |        ||:| | ||    |       ||
Db  2129  RLHRFAPPCCKPLLREEVSFRVGLHEYPVGSQLPC---EPE------PDVAVLTSM        2174

Qy  2053  L--PFRVDGHTPGVRMQLNLRDALETND-------CNSTNNTPSDEAAVSALVFK       2098
                    |  :|                         :||   |: |: |
Db  2175  LTDPSHITAEAAGRRLARGSPPSMASSSASQLSAPSLKATCTANHDSPDAELIEANLLWR   2234

Qy  2099  QEL------RRTNQLL-------EAISAGVDTTKLPAPSIEEVVVRKRQFRARTGSLTLP-  2145
          ||       :::                   |:|      | :
Db  2235  QEMGGNITRVESENKVVILDSFDPLVAEEDEREVSVPA--EILRKSRRFARALPVWARPD   2292

Qy  2146  ----------------PPPRSVP----------                            2152
                          |||:
Db  2293  YNPPLVETWKKKPDYEPPVVHGCPLPPPRSPPVPPPRRKKRTVVLTESTLSTALAELATKSF  2352

Qy  2153  ------GVSCPESLQRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTET  2206
                |: | :|     :|  ||| ||:|||            | :|       |
Db  2353  GSSSTSGITGDNTTTSSEP--APSGCPPDSDVESYSSMPPLEG--------E          2394

Qy  2207  GGGPDDLPSYPPKKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKK  2266
                   : : |                    | |
Db  2395  PGDPD-------LSDGSWSTVSS---------GADT------                    2414

Qy  2267  LGKSEFSCSMSYTWTDVISFKTASKVLS-ATRAITSGFLKQRSLVVYTEPRDAELRKQKV  2325
          | :: :|||||| ||  | |  |::|::   |  |: ||::::| : :|     ::|:||
Db  2415  --EDVVCCSMSYSWTGALVTPCAAEEQKLPINALSNSLLRHHNLVVSTTSRSACQRQKKV  2472

Qy  2326  TINRQPLFPPSYHKQVRLAKEKASKVVGVMNDYDEVAAHTPSKSAKSHITGLRGTDVRSG  2385
          | :|  :  | || |::     |||| | : : :||    | |||| :   | | ||
Db  2473  TFDRLQVLDSHYQDVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKF-GYGAKDVRC-  2530

Qy  2386  AARKAVLDL----QKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLLSYPHLEM  2441
          :||||| :     |:|   ::|||||||:| |||: |  :     |:|| |||  | :
Db  2531  HARKAVAHINSWKDLLEDSVTPI---DTTIMAKNEVFCVQPEKGGKRPARLLIVFPDLGV  2587

Qy  2442  RCVEKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMW--SPPAVGATCDTVCFDST  2498
```

```
Db   2588  RVCEKMALYDVVSKLPLAVMGSSYGFQYSPGQRVBFLVQAWKSKKTPMGFSYDTRCFDST 2647
                 ||   ||   :||:   |||   ::||:           .:  |  ||||:
Qy     87  LGILLDYPLGWIGDVTHTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLF 143

Qy   2499  ITPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGEMIAYDGREIGYRRCRSSGGVTT 2558
                 |||    ||::||  |:.|||  :    |||::||:||   ||:   ||||
Db    119  LGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLPGCGSFSIF 177

Db   2648  VTESDIRTEAIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGVLTT 2707
                 |:   || ||:|:       ||       |||:   ||| |||:|| |||:||
Qy    144  VVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC----AD 197

Qy   2559  SSSNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMG 2618
                 ::||:|  ||::| |  ||  |||| || ||::|  :  |||:|  || |
Db    178  LLALLSCLTVPASAYQVRN-SSGLYHVTNDCPNSSIVYEAADAILHTPGCVPCVREGNAS 236

Db   2708  SGGNTLTCYIKARAACRAAGLQDCTMLVCGDDLVVICESAGVQEDAASLRAFTEAMTRYS 2767
                 :|| |||:||  .:   || |  || ||||  |||: |||  ||:  ||| | :
Qy    198  ECWVPANPYISHSPNWTGTDSFPLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV--- 254

Qy   2619  APQDCVPQPKYSLEELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAA 2678
                 || |  |||  | :   |:|||||| |  :  :|||:||  :|||  |:|  |:|
Db    237  RCWVAVTPTVATRDGKLPTTQ-LRRHIDLLVGSATLCSALYVGDLCGSVFLVGQLFTFSP 295

Db   2768  APQGDPPQPEYDLEILTSCSSNVSVAHDCGAGKRVYYLTRDPTTPLARAAWETARHTPVNS 2827
                 ||  || |:|||::||||||||::|:   |   :||  :|||   ||   ||  | |
Qy    255  -RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGW--MAGKVEAVIFLTKLASQVPYAIA 310

Qy   2679  WIGYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPEVEDLPSIIAGVH 2738
                 |:  |:|| :  |     |:  |||:   :    : |   |   || | :|:|
Db    296  RRHW------TTDCNCSIYPGHITGHR--MAWDMMNWSPTAALVVAQLL-RIPQAIM 345

Db   2828  WLGNIIMFAPTLWARMILMTHFFSVLIARDQLEQALNCEIYGACYSIEPLDLPIIQRLH 2887
                 |||::: |   |  :::|  |||||:  ||:||   :||  || :|   :|  |: |
Qy    311  TMFSSVHYLAVGALIYYASRGKWYQLLLALMLY------IEATSGNPIRV- 354

Qy   2739  GIEAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL- 2797
                 |  |  |||| ||  |::| ::            |   |         |||
Db    346  DMIAGAHWGVLAGIAYFSMVGNWAKVLVVLLLFAGVDAETHVTGGNAGRTTAGLVGLLTP 405

Qy   2798  WHATSR----PLPLDKTSVARYTFENYCDVYSPEGDVFITPQRRLQKFPLVKYLAVIVFA 2853
                 |:   |          :||:|   ::  :                  :|   ||
Db    406  GAKQNIQLINTNGSWHINSTALNCNESLNTGWLAGLFYQHKFNSSGCPERLASCRRLTDF 465

Db   2854  LGLIAVGLAI 2863
                 |::|  |||
Qy    361  AEFCSPLMI--------PCPCHSYLSENV-SEVICYSPKWTRPITLEYNN- 401

Db   2948  WAVRTKLKLTPIAAAGRLDLSGWFTAGYSG-----GDIYHSVSHARPRW----- 2995
                 :|:  || | :   |   || :|:|  |||     ||||||:|:|||
Qy    402  ---SISWPYTIPGARGCMVKFKNNT------WGCCRIRNVPSY------CTMG---- 440

Db   2996  LLLLAAGVGI 3005
                 ||:|||||:|
Qy    466  AQGWGPISYANGSGLDERPYCWHYPPRPCGIVPAKSVCGPVYCFTPS---PVVVGTTDRS 522

Qy   2739  ...
Db   2888  GLSAFSLHSYSPGEINRVAACLRKLGVPPLRAWRHRARSVRARLLSRGGRAAICGKYLFN 2947
```

```
Query Match    20.9%;    Score 3214.5;    DB 14;    Length 3011;
Best Local Similarity 29.4%;  Pred. No. 5.6e-235;
Matches  956;    Conservative 440;    Mismatches 1213;    Indels 641;    Gaps 107;

Qy     8  TSPVPAPRTRKNKTQASYPVSIK----------------TSVERGQRAKRKVQRDARPR 51
              :|: :||:|:: |   :                  :|: | |  |||:||||
Db     3  TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR 59

Qy    52  NYK----IAGIHDGLQTLAQAALP------------------AHGWGRQDPRHKSRN 86
              ||:    |   |  | ||  | |                  || ||:|  | |||
Db    60  GRRQPIPKARRPEG-RTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRN 118
```

```
Db  2129 RLHRFAPPCKPLLREVSFRVGLHEYPVGSQLPC---EPE------------PDVAVLTSM 2174
Qy  2053 L--PFRVDGHTPGVRMQLNLRDALETND------------CNSTNNTPSDEAAVSALVFK 2098
Db  2175 LTDPSHITAEAAGRRLARGSPPSMASSASQLSAPSLKATCTANHDSPDAELIEANLLWR  2234
Qy  2099 QEL------RRTNQLL-----EAISAGVDTTKLPAPSIEVVRKRQFRARTGSLTLP-    2145
Db  2235 QEMGGNITRVESENKVVILDSFDPLVAEEDEREVSVPA--EILRKSRRFARALPWARPD  2292
Qy  2146 -----------PPPRSVP------                                    2152
Db  2293 YNPPLVETWKKKPDYEPPVVHGCPLPPPRSPPVPPPRKKRTVVLTESTLSTALAELATKSF 2352
Qy  2153 ---GVSCPESLQRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTET    2206
Db  2353 GSSSTSGITGDNTTTSSEP--APSGCPPPDSDVESYSSMMPLEG----------E     2394
Qy  2207 GGGPDDLPSYPPKKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKK 2266
Db  2395 PGDPD----LSDGSWSTVSS----GADT-                               2414
Qy  2267 LGKSEFSCSMSYTWTDVISFKTASKVLS-ATRAITSGFLKQRSLVYVTEPRDAELRKQKV 2325
Db  2415 --EDVVCCSMSYSWTGALVTPCAAEEQKLPINALSNSLLRHHNLVSTTSRSACQRQKKV  2472
Qy  2326 TINRQPLFPPSYHKQVRLAKEKASKVVGVWMDYDEVAAHTPSKSAKSHITGLRGTDVRSG 2385
Db  2473 TFDRLQVLDSHYQDVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKF-GYGAKDVRC- 2530
Qy  2386 AARKAVLDL----QKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEM 2441
Db  2531 HARKAVAHINSVWKDLLEDSVTPI---DTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGV 2587
Qy  2442 RCVERKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMW--SPDAVGATCDTVCFDST 2498
Db  2588 RVCERKMALYDVVSKLPLAVMGSSYGFQYSPGQRVEPLVQAWKSKKTPMGFSYDTRCFDST 2647
Qy  2499 ITPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTT  2558
Db  2648 VTESDIRTEEAIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGVLTT  2707
Qy  2559 SSSNSLTCWLKVNAAAEQAGMKNPRPLICGDDCTVIWKSAGADADKQAMRVFASWMKVMG  2618
Db  2708 SCGNTLITCYIKARAACRAAGLQDCTMLVCGDDLVVICESAGVQEDAASLRAFTEAMTRYS 2767
Qy  2619 APQDCVPQPKYSLEELTSCSSNVTSGITKSGKPYYFLTRDRPILPGRCSAEGLGYNPSAA  2678
Db  2768 APPGDPPQPEYDLELJTSCSSNVSVAHDGAGKRVYLTRDPTTPLARAAWETARHTPVNS   2827
Qy  2679 WIGYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWGKNYTVPVEDLPSIIAGVH   2738
Db  2828 WLGNIIMFAPTLWARMILMTHFFSVLIARDLEQALNCEIYGACYSIEPLDLPPIIQRLH   2887
Qy  2739 GIEAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL-  2797
Db  2888 GLSAFSLHSYSPGEINRVAACLRKLGVPPLRAWRHRARSVRARLLSRGGRAAICGKYLFN  2947
Qy  2798 WHAT5R----PLPDLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKLVAVIVFA  2853
Db  2948 WAVRTKLKLTPIAAAGRLDLLSGWFTAGYSG-----GDIYHSVSHARPRW           2995
Qy  2854 LGLIAVGLAI 2863
Db  2996 LLLLAAGVGI 3005
```

```
Db  1077 VCWTVYHGAGTRTIASPKGPV-..IQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVT 1134
Qy  1049 RLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAA---RNSGGSVS  1105
Db  1135 RHADVIPVRRRGDSRGSLLSPRFISYLKGSSGGPLLCPAGHAVGLFRAAVCTRGVAKAVD  1194
Qy  1106 QIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQEKYEV  1165
Db  1195 FIPVENLGTTMRSPVFTDNS---SPPAVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKV  1251
Qy  1166 LVLNPSVATTASMPKYMHATYGVNPNCYFNGKCNTGASLTYSTYGMYLT-GACSRN-YD   1223
Db  1252 LVLNPSVAATLGFGAYMSKAHGVDDNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYD  1311
Qy  1224 VIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDEG  1283
Db  1312 IIICDECHSTDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVSHPNIEEVALSTTG  1371
Qy  1284 TIPFHGKKIKEENLKKGRHLIFEATKKGHCDELANELARKGITAVSYYRGCDISKIP-EGD 1342
Db  1372 EIPFYGKAIPLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGD  1431
Qy  1343 CVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPTFTMGVRVCGVSAIVKGQRRGRT   1402
Db  1432 VVVSTDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTLPQDAVSRTQRRGRT    1491
Qy  1403 GRGRAGIYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAI   1462
Db  1492 GRGKPGIYRFVAPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVC  1551
Qy  1463 GANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLQLCHQYGYAAPNDAPRWQG-  1520
Db  1552 QDHLEFWEGVFTGLTHIDAHFLSQTKQSGENFPYLVAYQATVCARAQAPPSWDQMWKCL   1611
Qy  1521 ARLGKK---PCGVLWRLDGADACPGPEPSEV------TRYQM-CFT---EVNTSGTAALA  1567
Db  1612 IRLKPTLHGPTPLLYRL-------GAVQNEVTLTHPITKYIMTCMSADLEVVTS-TWVLV  1663
Qy  1568 VGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI------VEECASFIPL   1621
Db  1664 GGVLAALAAYCLSTGCVVIVGR-----IVLSGKPAIIPDREVLYQEFDEMEECSQHLPY   1717
Qy  1622 --EAMVAA----IDKLKSTITTTSPF----TLETALEKLNTFLGPHAARTILAIIEYCCG  1670
Db  1718 IEQGMMLAEQFKQKALGLLQTASRHAEVITPAVQTNWQKLEVFWAKHMWNFISGIQYLAG  1777
Qy  1671 LVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFPGGAIASKLTDARGALAFMWAGAAGT 1730
Db  1778 LSTLPGNPAIASLMAFTAAVTSPLTTGQTLFNILGGWVAAQLAAPGAATAFVGAGLAGA   1837
Qy  1731 ALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVVSAFNPAAGVV  1787
Db  1838 AIG---SVGLGKVLVDILAGYGAGVAGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVV  1894
Qy  1788 GVLSACAMFALTTAGPD----HWPNRLLITMLARSNTVCNEYFIATRDIRRKILGILEAST 1843
Db  1895 GVV--CAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAARVTAILSSLT  1952
Qy  1844 PWSVISACIRWLHTPTEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFY  1903
Db  1953 VTQLLRRLHQWISSECTTPCS-GSWLRDIWDWICEVLSDFKTWLKA--KLMPQLPGIPFV  2009
Qy  1904 SCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLLYKGPRTCSNYWRGAVPVNARLCG 1963
Db  2010 SCQRGYRGVWRGDGIMHTRCHCGAEITGHVKNGIMRIV-GPRTCRNMWSGTFPINAYTTG  2068
Qy  1964 SARPDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFTQVP--------    2010
Db  2069 PCTPLPAPNYKFALWRVSAEEYVEIRRVGDFHYVSGMTTDNLKCPCQIPSPEFFTELDGV  2128
Qy  2011 ---------PTLRAAVAVD------GVQVQCYLGEPKTPWTTSACCYGPDGKGKTVK     2052
```

```
; APPLICANT: SANGAR, DAVID V.
; APPLICANT: LEMON, STANLEY M.
; TITLE OF INVENTION: Chimeric GB Virus B (GBV-B)
; FILE REFERENCE: UTSG:258US
; CURRENT APPLICATION NUMBER: US/10/189,359
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: .10/189,359
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 14
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-189-359-14

Query Match        20.9%;  Score 3214.5;  DB 15;  Length 3011;
Best Local Similarity  29.4%;  Pred. No: 5.6e-235;
Matches 956;  Conservative 440;  Mismatches 1213;  Indels 641;  Gaps 107;

Qy      8  TSPVPAPRTRKNKTQASYPVSIK----------------------TSVERGQRAKRKVQRDARPR    51
           |:|  |::|::| |  :| |                      :|||||
Db      3  TNPKPQRKTKRNTNRR---PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR    59

Qy     52  NYK-----IAGIHDGLQTLAQAALP----------------------AHGWGRQDPRHKSRN    86
           |||      ||||| | | |:  ||                      |||  ||| | |
Db     60  GRRQPIPKARRPEG-RTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRN   118

Qy     87  LGILLDYPLGWIGDVTTHTPLVGAPGYPWPLYGVPLVAGAVVRPVCQIVRLLEDGVNWATG---WFGVHLF   143
           |||:| || |  || || |||| |  || | | | |   :  | ||:||  ||:|| |||     |||||
Db    119  LGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLPGCSFSIF   177

Qy    144  VVCLLS-LACPCSGARVTDPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC------AD    197
           ::| ||  | |||   | ||| |||| :|:| :|:  | | |||  ||  |         |
Db    178  LLALLSCLTVPASAYQVRN-SSGLYHVTNDCPNSSIVEAADAILHTPGCVPCVREGNAS   236

Qy    198  ECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV---   254
           ||| || |: ||: ||:|:  :     |||:  |  |    :|:|:  | |  :|
Db    237  RCWVAVTPTVATRDGKLPTTQ-LRRHIDLLVGSATLCSALYVGDLCGGSVFLVGQLFTSP   295

Qy    255  -RHWLIHIDLNETGTCYLEVPTGIDPGFIGFIGW---MAGKVEAVIFLTKLASQVPYAIA   310
           |||| : |||| ||| |:|| |  | |  |:      |||:|:  |:||||:  | |
Db    296  RRHW------TTQDCNCSIYPGHITGHR--MAWDMMWNWSPTAALVVAQLL-RIPQAIM   345

Qy    311  TMFSSVHYLAVGALIYYASRGKWYQLLLALMLY------IEATSGNPIRV-----   354
           | ::|||  | | |:| |:  |  |||| |||          |:|||| |
Db    346  DMIAGAHWGVLAGIAYFSMVGNWAKVLVVLLLFAGVDAETHVTGGNAGRTTAGLVGLLTP   405

Qy    355  -----------------------------PTGC----------SI   360
                                        |||           :|
Db    406  GAKQNIQLINTNGSWHINSTALNCNESLNTGWLAGLFYQHKFNSSGCPERLASCRRLTDF   465

Qy    361  PCPCHSYLSENV-SEVICYSPKWTRPITLEYNN--   401
           |||  | :  |  | ||: | |    ||| | |
Db    466  AQGWGPISYANGSGLDERPYCWHYPPRPCGIVPAKSVCGPVICFTPS---PVVVGTTDRS   522

Qy    402  ---SISWYPTIPGARGCMVKFKNNT-------WGCCRIRNVPSY-------CTMG--   440
               :|  |  | ||||:| |                | |  |           |
Db    523  GAPTYSW------GANDTDVFVLNNTRPPLGNWFGCTWMNSTGFTKVCGAPPCVIGGVGN   576

Qy    441  ------TDAVWNDTRNTYEACGVTPWLT-------------TAWHNGSALKLAILQ---YPGSK   482
                    |  | || :  |:|  | |             || | | || :|:|    ||||
Db    577  NTLLCPTDCFRKHPEATYSRCGSGPWITPRCMVDYPYRLWHYPCTINYTIFKVRMYVGGV   636

Qy    483  E--MFKPHNWMSG-HLYFEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVP   539
           |   | |     | |||  |      | |  :  |         |  || |  | :
Db    637  EHRLEAACNWTRGERCDLEDRDR------SELSPLLLSTTQWQVLPCS---------   678

Qy    540  QGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILLGCGSKYILLA   599
           |  || | :    ||  | | :|  :|: | | | | ::  | ||:| |  | :|:|
Db    679  --FTTLPALSTGLIHLHQNIVDVQYLYGVGSSIASWAIKWEYVVLFLLLADARVCSCL   735

Qy    600  YLCYLSLCFGRASGYPLRPVLFSQSYLQAGWDVLSKAQVAPPALIFFICCYLRCRLRYAA   659
           | : |  |                :|:     :  |   | | | |  ||  |  || |
Db    736  WMMLL-|                                          ISQEAALENLVI-   758
```
```
Qy    660  LLGFVPMAAGLP--LTFFVAAAAAOPDYDWVRLLVAGLVLWAGRNRGHRIALLVGPWPL   717
           |||  |  ||    | || |      |  : |   | |  | |:    ||||| ||||
Db    759  LAG---THGLVSFLVFFCFA---------------WYLK------GRWVPGAVYALYGMWPL   796

Qy    718  VALLTLLHLVTPASAFDTEI---IGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYLWQ   774
           |||| ||:| | |:  |       |  | | | | ::|   : | | | |  | || | |
Db    797  -LLLLALPQRAYALDTEVAASCGGVVL-VGLMALTLSPYYKRYISWCMWHLQYFLT   851

Qy    775  RWEN----WFWNVTLRPERPFLVLVCFPGATYDALVTFCVCHVALLC-LTSSAASPFGT-   828
           | |      | |  | |  |||:|| |||  ||||||| |||| | | | : | |||
Db    852  RVEAQLHVWVPPLNVRGGRDAVILL-------MCVVHPTLVFDITKLLLAIFGPL   899

Qy    829  -----------DSRVRAHRMLVRLGKCHAWYSHYVLKFFLLVFGENGVFFYKHLH--GD   874
                      |::| |: |:|:|| | |: |||: |  ||:| |   || | :  ||
Db    900  WILQASLLKVPYFVVRVQGLLRICALARKIAG-GHVVQMAIIKLGALTGTVYNHLTPLRD   958

Qy    875  VLPNDFASKJPLQEP--FPPFEGKARVYRNEGRRLACGDTVDGLPVVARLGDLVPAGLA-   931
           | |    |      |  | |  :  |: | :  | |:|| : || | | | | : | |
Db    959  WAHNGLRDLAVAVEPVVFSRMETKLITWGAD--TAACGDIINGLPVSARRGQEILLGPAD   1016

Qy    932  -MPPDGWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSYMGFVCDN   990
            ||  ||  ||  |  :| | || | | |   ||||||:  | || ::  | |||:
Db    1017 GMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTATQTFLATCING   1076

Qy    991  VLYTAHHGSKGRRLAHPTGSIHPITVDAANDQI--YQPPCGAGSLTRCSCGETKGYLVT   1048
           |  | | |   |||: | |:  | ||   |:|   :| |   :| | | | |  |:||
Db    1077 VCWTVYHGAGTRTTIASPKGPV-IQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVT   1134

Qy    1049 RLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAA----RNSGGSVS   1105
           |    |  | | ||| |||| |  ||  |  ||:|| |:| |:|         ||  | |
Db    1135 RHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAVCTRGVAKAVD   1194

Qy    1106 QIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQEKYEV   1165
           |   ||||    | | |||: | ||| |   :|  ||:||||||||||| |  | :::|
Db    1195 FIPVENLGTTMRSPVFTDNS---SPPAVQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKV   1251

Qy    1166 LVLNPSVATTASMPKYMHATGYVNPNCYFNGKCTNTGASLTYSYGMYLT-GACSRN-YD   1223
           ||||| |||| ||    | |||:|||:  ::|| ||||::|:|:||:   |||: ||
Db    1252 LVLNPSVAATLGFGAYMSKAHGVDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYD   1311

Qy    1224 VIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANITEIQLTDEG   1283
           :||||||| |||| |:|||| || |: ||:|| |||||||||: ||:|| : | :| |
Db    1312 IIICDECHSTDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVSHPNIEEVALSTTG   1371

Qy    1284 TIPPHGKKIKEENLKKGRHLIFEATKGHCDELANELARKGITAVSYRGCDISKIP-EGD   1342
           |||  |: || |||  ||||||| |: | ||| ||  |  | |||| ||||:|  :||
Db    1372 EIPFYGKAIPLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGD   1431

Qy    1343 CVVVATDALCTGYTGDFDSVYDGSLMVEGTCHVDLDPFTMGVRVCGVSAIVKGQRRGRT   1402
           |||: || |||||||||||||| |:||| |||||||||||:::| ||  | |||||||
Db    1432 VVVVSTDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTLPQDAVSRTORRGRT   1491

Qy    1403 GRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAKAWYGLSSTEAQTILDTYRTQPGLPAI   1462
           ||||:|||: |||||| |||| ||| |||||||  ||  ||:| ::|||||| | ||
Db    1492 GRGKPGIYRFVAPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTGLPVC   1551

Qy    1463 GANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLQLCHOYGYAAPNDAPRWQG-   1520
           | :| |   |   :|| |  | | |:|| |: |:| | |:| |:|| |: |  | |
Db    1552 QDHLBFWEGVFTGLTHIDAHFLSQTKQSGENFPYLVAYQATVCARAQAPPPSWDQMWKCL   1611

Qy    1521 ARLGKK------PCGVLWRLDGADACPGPEPSEV------TRXQM-CFT--EVNTSGTAALA   1567
           ||| :|      |      || |: :|  | | |      |: :|      |:  | | |
Db    1612 IRLKPTLHGPTPLLYRL-------GAVQNEVTLTHPITKYIMTCMSADLEVVTS-TWVLV   1663

Qy    1568 VGVGAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEI------VEECASFIPL   1621
           :| | |:  | ||| | | || | :| ::| ||: |:||::       || ||:
Db    1664 GGVLAALAAYCLSTGCVVIVGR------IVLSGKPAIIPDREVLYQEFDEMEECSQHLPY   1717

Qy    1622 --EAMVAA---IDKLKSTITTTSPP-------TLETALEKLNTFLGPHAATILAIIEYCCG   1670
             : |         :|  |  | |        |   |  : :  ||  |    |   | |
Db    1670 YLCYLSLCFGRASGYPLRPVLFSQSYLQAGWDVLSKAQVAPPALIFFICCYLRCRLRYAA   659
```

```
Db  1718  IEQMMLAEQFKQKALGLLQTASRHAEVITPAVQTNWQKLEVFWAKHMWNFISGIQYLAG 1777
Qy  1671  LVTLPDNPFAASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGT 1730
Db  1778  LSTLPGNPAIASLMAFTAAVTSPLTTGQTLLFNILGGWVAAQLAAPGAATAFVGAGLAGA 1837
Qy  1731  ALGTWTSVGF---VFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAAGVV 1787
Db  1838  AIG--SVGLGKVLVDILAGYGAGVAGALLVAFKIMSGEVPSTEDLVNLLPAILSPGALVV 1894
Qy  1788  GVLSACAMFALTTAGPD---HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEAST 1843
Db  1895  GVV--CAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAARVTAILSSLT 1952
Qy  1844  PWSVISACIRWLHITPTEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFY 1903
Db  1953  VTQLLRRLHQWISSECTTPCS-GSWLRDIWDWICEVLSDFKTWLKA--KLMPQLPGIPFV 2009
Qy  1904  SCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCG 1963
Db  2010  SCQRGVRGVWRDGIMHTRCHCGAEITGHVKNGTMRIV-GPRTCRNMWSGTFPINAYTTG 2068
Qy  1964  SARPDT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFTQVP--------- 2010
Db  2069  PCTPLPAPNYKFALMVRVSAEEYVEIRRVGDFHYVSGMTTDNLKCPCQIPSPEFFTELDGV 2128
Qy  2011  --------PTLRAAVAVD-------GVQVQCYLGEPKTPWTTSACCYGPDGKGKTVK 2052
Db  2129  RLHRFAPPCCKPLLREEVSFRVGLHEYPVGSQLPC---EPE--------PDVAVLTSM 2174
Qy  2053  L--PFRVDGHTPGVRMQLNRDALETND--------------CNSTNNTPSDEAAVSALVFK 2098
Db  2175  LTDPSHITAEAAGRRLARGSPPSMASSSASQLSAPSLKATCTANHDSPDAELIEANLLWR 2234
Qy  2099  QEL--------RRTNQLL------EAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLP- 2145
Db  2235  QEMGGNITRVESENKVVILDSFDPLVAEBDEREVSVPA--EILRKSRRFARALPVWARPD 2292
Qy  2146  ---------------------------PPPRSVP----------------------- 2152
Db  2293  YNPPLVETWKKPDYEPPVVHGCPLPPPRSPPVPPPRKKRTVVLTESTLSTALAELATKSF 2352
Qy  2153  ------GVSCPESLQRSDPLEGPSNLPPSPPVLQLAMPMDLLGAGECNPPTAIGCAMTET 2206
Db  2353  GSSTSGITGDNTTTSSEP--APSGCPPDSDVESYSSMPPLEG---------------E 2394
Qy  2207  GGGPDDLPSYPPKKEVSEWSDESWSTATTASSYVTGPPYYPKIRGKDSTQSAPAKRPTKKK 2266
Db  2395  PGDPD----------------LSDGSWSTVSS--------------GADT--------- 2414
Qy  2267  LGKSEFSCSMSYTWTDVISFKTASKVLS-ATRAITSGFLKQRSLVYVTEPRDAELRKQKV 2325
Db  2415  --EDVVCCSMSYSWTGALVTPCAAEEQKLPINALSNSLLRHHNLVYSTTSRSACQRQKKV 2472
Qy  2326  TINRQPLFPPSYHKQVRLAKEKASKVVGVNWDYDEVAAHTPSKSAKSHITGLRGTDVRSG 2385
Db  2473  TFDRLQVLDSHYQDVLKEVKVAAASKVKANLLSVEEACSLTPPHSAKSKF-GYGAKDVRC- 2530
Qy  2386  AARKAVLDL-----QKCVERAGEIPSHYRQTVIVPKEEVVKTPQKPTKKPPRLLISYPHLEM 2441
Db  2531  HARKAVAHINSVWKDLLLEDSVTPI---DTTIMAKNEVFCVQPEKKGRKPARLIVFPDLGV 2587
Qy  2442  RCVEKMYYGGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMW--SPDAVGATCDTVCFDST 2498
Db  2588  RVCEKMALVDVVSKLPLAVMGSSYGFQYSPGQRVEFLVQAWKSKKTPMGFSYDTRCFDST 2647
Qy  2499  ITPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTT 2558
Db  2648  VTESDIRTEEAIYQCCDLDPQARVAIKSLTERLYVGGPLITNSRGENCGYRRCRASGVLTT 2707
Qy  2559  SSSNSLTCWLKVNAAAEAQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMG 2618
Db  2708  SCGNTLTCYIKARAACRAAGLQDCTMLVCGDDLVVICESAGVQEDAASLRAFTEAMTRYS 2767
```

```
Qy  2619  APQCVPQPKYSLEELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAA 2678
Db  2768  APPGDPPQPEYDLELITSCSSNVSVAHDGAGKRVYLTRDPTTPLARAAWETARHTPVNS 2827
Qy  2679  WIGYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVH 2738
Db  2828  WLGNIIMFAPTLWARMILMTHFFSVLIARDQLEQALNCEIYGACYSIEPLDLPPIIQRLH 2887
Qy  2739  GIEAFSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL- 2797
Db  2888  GLSAFSLHSYSPGEINRVAACLRKLGVPPLRAWRHRARSVRARLLSRGGRAAICGKYLFN 2947
Qy  2798  WHATSR----PLPDLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFA 2853
Db  2948  WAVRTKLKLTPIAAAGRLDLSGWFTAGYSG----GDIYHSVSHARPRW---------FWFC 2995
Qy  2854  LGLIAVGLAI 2863
Db  2996  LLLLAAGVGI 3005
```

RESULT 15

US-11-006-313-20

```
  Query Match             20.9%;  Score 3214.5;  Pred. No. 5.6e-235;
  Best Local Similarity   29.4%;
  Matches  956;  Conservative  440;  Mismatches 1213;  Indels  641;  Gaps  107;
```

```
Qy     8  TSPVPAPRTRKNKQTQASYPVSIK--------------------TSVERGQRAKRKVQRDARPR  51
Db     3  TNPKPQRKTKRNTNRR----PQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPR  59
Qy    52  NYK----IAGIHDGLQTLAQAALP-----------------AHGWGRQDPRHKSRN  86
Db    60  GRRQPIPKARRPEG-RTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRN 118
Qy    87  LGILLDYPLGWIGDVTTHTPLVGPLVAGAVVRPVCCQIVRLLEDGVNWATG---WFGVHLF 143
Db   119  LGKVIDTLTCGFADLMGYIPLVGAPLGGA-ARALAHGVRVLEDGVNYATGNLPGCSFSIF 177
Qy   144  VVCLLS-LACPCSGARVTDPPDNTTILTNCCQRNQVIYCSPSTCLHEPGCVIC-----AD 197
Db   178  LLALLSCLTVPASAYQVRN-SSGLYHVTNDCPNSSIVEAADAILHTPGCVPCVREGNAS 236
```

```
Qy   198  ECWVPANPYISHPSNWTGTDSFLADHIDFVMGALVTCDALDIGELCGACVLVGDWLV---  254
          ||||    :.:        ||| |  :.:|||.  :||||  .::.|||||
Db   237  RCWVAVTPTVATRDGKLPTTQ-LRRHIDLLVGSATLCSALYVGDLCGSVFLVGQLFTPSP  295

Qy   255  -RHWLIHIDLNETGTCYLEVPTGIDPGFLGFIGW---MAGKVEAVIFLTKLASQVPYAIA  310
          :||| |   |  |.   :  ||  : .|     |.  ||  :|     :    ::.|| :
Db   296  RRHW-------TTQDCNCSIYPGHITGHR--MAWDMMNWSPTAALVVAQLL-RIPQAIM   345

Qy   311  TMFSSVHYLAVGALIYYASRGKWYQLLLALMLY----IEATSGNPIRV---------   354
          .:  |.: |:  | || ..  |  ||| |  |       ||.|
Db   346  DMIAGAHWGVLAGIAYFSMVGNWAKVLVVLLLFAGVDAETHVTCGNAGRTTAGLVGLLTP  405

Qy   355  ------------------------PTGC--------SI   360
                                   :||       ||
Db   406  GAKQNIQLINTNGSWHINSTALNCNESLNTGWLAGLFYQHKFNSSGCPERLASCRRLTDF  465

Qy   361  AEFCSPLMI----------------PCPCHSYLSENV-SEVICYSPKWTRPITLEYNN--  401
          | |          |         ||.  |    ||  ||:      |   ||| |
Db   466  AQGWGPISYANGSGLDERPYCWHYPPRPCGIVPAKSVCGPVYCFTPS---PVVVGTTDRS  522

Qy   402  ---SISWYPYTIPGARGCMVKFKNNT------WGCCRIRNVPSY-------CTMG----  440
             :|. .|      | |            .|||    |          |.
Db   523  -----GANDTDVFVLNNTRPPLGNWFGCTWMNSTGFTKVCGAPPCVIGGVGN        576

Qy   441  ------TDAVWNDTRNTYEACGVTPWLT-----------TAWHNGSALKLAILQ---YPGSK  482
                ||.|  |   |.| ||||              ||:|.|:|  |:|     .||.
Db   577  NTLLCPTDCFRKHPEATYSRCGSGSPWIITPRCMVDYPYRLWHYPCTINYTIFKVRMYVGGV  636

Qy   483  E--MFKPHNWMSG-HLYPEGSDTPIVYFYDPVNSTLLPPERWARLPGTPPVVRGSWLQVP  539
          :  ::|.      |  |   ||  :.   |    .|  |  |||     |  |||  |
Db   637  EHRLEAACNWTRGERCDLEDRDR------SELSPLLLSTTQWQVLPCS-            678

Qy   540  QGFYSDVKDLATGLITKDKAWKNYQVLYSATGALSLTGVTTKAVVLILGLCGSKYLILA   599
          |||.||| |  ||:||| |     :  |  |.|:||:|||  |:.:|| |||||:|
Db   679  ---FTTLPALSTGLIHLHQNIVDVQYLYGVGSSIASWAIKWEYVVLFLLLADARVCSCL   735

Qy   600  YLCYLSLCFGRASGYPLRPVLPSQSYLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAA  659
          |   |.|    ||   |      :  :.|    | .|       |    ||   .:
Db   736  WMML-------------------ISQAEAALENLVI------------LNAAS        758

Qy   660  LLGFVPMAAGLP--LTFFVAAAAAQPDYDWVRLLVAGLVLWAGRNRGHRIALLVGPWPL   717
           |||||  ||   :|   |||||||  : |  |    |  ||    ||     | |
Db   759  LAG----THGLVSFLVFFCFA--------WYLK---------GRWVPGAVVALYGMWPL   796

Qy   718  VALLTLLHLVTPASAFDTEI--IGGLTIPPVVALVVMSRFGFFAHLLPRCALVNSYIWQ   774
          ||| ::||:|| ||.| |   |||  |||||:.||.| |    |.| :  |: ||
Db   797  --LLLLALPQRAYALDTEVAASCGGVVL---VGLMAILTLSPYYKRYISWCMWWLQYFLT  851

Qy   775  RWEN-----WFWNVTLRPERFFLVLVCFPGATYDALVITFCVCHVALLC-LTSSAASFFGT-  828
          | :|     |.|  |   .  :      |     |  :||  | :|||   :|:.||.
Db   852  RVEAQLHVWVPPLNVRGGRDAVILL---------MCVVHPTLVFDITKLLLAIFGPL     899

Qy   829  ----------------DSRVRAHRMLVRLGKCHAWYSHYVLKFFLLVFGENGVFFYKHLH--GD  874
                          :|  ||.||||.|:|     ||  ||.||  ||.| |||| |    .|
Db   900  WILQASLLKVPYFVRVQGLLRICALARKIAG-GHYVQMAIIKLGALTGTYVNHLTPLRD       958

Qy   875  VLPNDFASKLPLQEP--FFPPEGKARVVRNEGRRLACGDTVDGLPVVARLGDLVFAGLA-  931
          ||| |  |   | |  ||||.|  |||.  |  ||.||| :.|| |:.||.:||:
Db   959  WAHNGLRDLAVAVEPVVFSRMETKLITWGAD--TAACGDIINGLPVSARRGQEILLGPAD  1016

Qy   932  -MPPDGWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGTIFRLGSLATSYMGFVCDN  990
            |  |   ||  | |:||||.|.||.||||||.||||||||||||||||||||||:|.
Db  1017  GMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTATQTFLATCING  1076

Qy   991  VLYTAHHGSKGRRLAHPTGSIHPITVDAANDQDI--YQPPCGAGSLTRCSCGETKGYLVT  1048
          |||.  .|    | |   |  | :|| |    :    ||||| |||.| |||||||:||
Db  1077  VCWTVYHGAGTRTTIASPKGPV--IQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVT  1134

Qy  1049  RLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSGHVIGMFTAA---RNSGGGSVS  1105
          ||:||:|  .. | | |||  ||:|| |.||||:||||||||||.||    |:||| ||
Db  1135  RHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAVCTRGVAKAVD  1194

Qy  1106  QIRVRPLVCAGYHPQYTAHATLDTKPTVPNEYSVQILIAPTGSGKSTKLPLSYMQEKYEV  1165

Qy  1195  FIPVENLGTTMRSPVFTDNS---SPPAVQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKV  1251
          ||||       |          |  :.:|   :  |      ||||||||||||     :.:.|
Db  1166  LVLNPSVATTASMPKYMHATYGVNPNCYFNGKCNTNGASLTYSYGMYLT-GACSRN-YD  1223

Qy  1252  LVLNPSVAATLGFGAYMSKAHGVDPNIRTGVRTITTGSPITYSTVGKFLADGGCSGGAYD  1311
Db  1224  VIICDECHATDATTVLGIGKVLTEAPSKNVRLVVLATATPPGVIPTPHANITEIOLTDEG  1283

Qy  1312  IIICDECHSTDATSILGTVLDQAETAGARLVVLATATPPGSVTVSHPNIEEVALSTTG   1371
Db  1284  TIPFHGKKIKBENLKKGRHLIFEATKGHCDELANELARKGITAVSYYRGCDISKIP-EGD  1342

Qy  1372  EIPFYGKAIPLEVIKGGRHLIFCHSKKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGD  1431
Db  1343  CVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLDPFTMGVRVCGVSAIVKGQRRGRT   1402

Qy  1432  VVVVSTDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTLPQDAVSRTQRRGRT.  1491
Db  1403  GRGRAGIYYYVDGSCTPSGMVPECNIVEAPDAAKAWYGLSSTEAQTILDTYRTQPGLPAI  1462

Qy  1492  GRGKPGIYRFVAPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVC  1551
Db  1463  GANLDEWADLFS-MVNPEPSFVNTAKRTADNYVLLTAAQLCHOYGYAAPNDAPRWQG-   1520

Qy  1552  QDHLEFWEGVFTGLTHIDAHFLSQTKQSGENFPYLVAYQATVCARAQAPPSWDQMWKCL  1611
Db  1521  ARLGKK------PCGVLWRLDGADACPGPEPSEV------TRYQM-CFT---EVNTSGTAALA  1567

Qy  1612  IRLKPTLHGPTPLLYRL-----GAVQNEVTLTHPITKYIMTCMSADLEVVTS-TWVLV   1663
Db  1568  VGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEEEI------VEECASFIPL  1621

Qy  1664  GGVLAALAAVCLSTGCVVIVGR-------IVLSGKPAIIPDREVLYQEFDEMEECSQHLPY  1717
Db  1622  --EAMVAA---IDKLKSTITTTSPP-------TLETALEKLNTFLGPHAATLAIIEYCCG  1670

Qy  1718  IEQGMMLAEQFKQKALGLLQTASRHAEVITPAVQTNWQKLEVFWAKHMWNPISGIQYLAG  1777
Db  1671  LVTLPDNPFASCVFAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAGAAGT  1730

Qy  1778  LSTLPGNPAIASLMAFTAAVTSPLTTGQTLLFNILGGWVAAQLAAPGAATAFVGAGLAGA  1837
Db  1731  ALGTWTSVGF---VFDMLCGYAAASSTACLTFKCLMGEWPTMDQLAGLVVSAFNPAAGVV  1787

Qy  1838  AIG---SVGLGKVLVDILAGYGAGVAGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVV  1894
Db  1788  GVLSACAMFALTTAGPD---HWPNRLLTMLARSNTVCNEYFIATRDIRRKILGILEAST   1843

Qy  1895  GVV--CAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPEBSDAAAARVTAILSSLT  1952
Db  1844  PWSVISACIRWLHTPTEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGCPFY  1903

Qy  1953  VTQLLRRLHQWISSECTTPCS-GSWLRDIWDWICEVLSDFKTWLKA--KLMPQLPGIPFV  2009
Db  1904  SCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVNARLCG  1963

Qy  2010  SCQRGYRGVWRGDGIMHTRCHCGAEITGHVKNGTMRIV-GPRTCRNMWSGTFPINAYTTG  2068
Db  1964  SARPDPT-DWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPNV-CFTQVP----------  2010

Qy  2069  PCTPLPAPNYKFALMRVSABEYVEIRRVGDFHVVSGMTTDNLKCPCQIFSPEFFTELDGV  2128
Db  2011  --------------GVQVQCYLGBPKTPWTTSACCYGPDGKGKTVK               2052

Qy  2129  RLHRRAPPCKPLLREEVSFRVGLHEYPVGSQLPC---EPB---------PDVAVLTSM   2174
Db  2053  L--PFRVDGHTPGVRMQLNLRDALETND----CNSTNNTPSDEAAVSALVFK         2098

Qy  2175  LTDPSHITAEAAGRRLARGSPPSMASSSASQLSAPSLKATCTANHDSPDAELIEANLLWR  2234
Db  2099  QEL------RRTNQLL-----EAISAGVDTTKLPAPSIEEVVRKRQFRARTGSLTLP-   2145
```

```
Db   2235 QEMGGNITRVESENKVVILDSFDPLVAEEDEREVSVPA--EILRKSRRFARALPVWARPD 2292
                                        ||||||
Qy   2146 -----------------------------PPPRSVP------------------------- 2152

Db   2293 YNPPLVETWKKPDYEPPVVHGCPLPPPRSPPVPPPRKKRTVVLTESTLSTALAELATKSF 2352
                :  :  :         ||||| :    : :  ||||
Qy   2153 ------GVSCPESLQRSDPLEGPSNLPPSPPVLQLAMPMPLLGAGECNPFTAIGCAMTET 2206

Db   2353 GSSSTSGITGDNTTTSSEP--APSGCPPDSDVESYSSMPPLEG--------------E   2394
           |::  : |:|                 ::      : |||
Qy   2207 GGGPDDLPSYPPKKEVSEWSDESWSTATTASSYVTGPPYPKIRGKDSTQSAPAKRPTKKK 2266

Db   2395 PGDPD-----------LSDGSWSTVSS---------------GADT-------------  2414
           |:|||                   ::                   ::
Qy   2267 LGKSEFSCSMSYTWTDVISFKTASKVLS-ATRAITSGFLKQRSLVYVTEPRDAELRKQKV 2325

Db   2415 --EDVVCCSMSYSWTGALVTPCAAEEQKLPINALSNSLLRHHNLVYSTTSRSACQRQKKV 2472
             : ||||||||| |||      ||||    |  :||    ||||     :  ||||||
Qy   2326 TINRQPLFPPSYHKQVRLAKEKASKVVGVMWDYDEVAAHTPSKSAKSHITGLRGTDVRSG 2385

Db   2473 TFDRLQVLDSHYQDVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKF-GYGAKDVRC-  2530
           :  |  |    |  | :| | ||:|||         :  ||| ||:|||  |||| |||
Qy   2386 AARKAVLDL-----QKCVEAGEIPSHYRQTVIVPKEEVFVKTPQKPTKKPPRLISYPHLEM 2441

Db   2531 HARKAVAHINSVWKDLLEDSVTPI---DTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGV 2587
            ||||||      :|  |||   ||     :: ||||||    :| |  |||||: ||||
Qy   2442 RCVEKMYYGQVAPDVVKAVMGDAYGF-VDPRTRVKRLLSMW--SPDAVGATCDTVCFDST 2498

Db   2588 RVCEKMALYDVVSKLPLAVMGSSYGFQYSPGQRVEFLVQAWKSKKTPMGFSYDTRCFDST 2647
            |||||      |  |||||||    | ::   ||  | || |       ||  ||||||
Qy   2499 ITPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMIAYDGREIGYRRCRSSGVYTT 2558

Db   2648 VTESDIRTEAIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGVLTT  2707
           :|  || |  :|   :| | || ||:|| ||||| |: : |||| |||||||:||| ||
Qy   2559 SSSNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWKSAGADADKQAMRVFASWMKVMG 2618

Db   2708 SCGNTLTCYIKARAACRAAGLQDCTMLVCGDDLVVICESAGVQEDAASLRAFTEAMTRYS 2767
           | ||||||: ||  ||||| ||| |::|||| :||| |||  ||||||| ||  ||||
Qy   2619 APQDCVPQPKYSLELTSCSSNVTSGITKSGKPYYFLTRDPRIPLGRCSAEGLGYNPSAA  2678

Db   2768 APPGDPPQPEYDLELITSCSSNVSVAHDGAGKRVYLTRDPTTPLARAAWETARHTPVNS  2827
           || | ||||| ||| :|||||| :  : ||||||||||||| ||||  || || |||:
Qy   2679 WIGYLIHHYPCLWVSRVLAVHFMEQMLFEDKLPETVTFDWYGKNYTVPVEDLPSIIAGVH 2738

Db   2828 WLGNIIMFAPTLWARMILMTHFFSVLIARDQLEQALNCEIYGACYSIEPLDLPPIIQRLH 2887
           | | :: :  |||| :::  ||  :| | || |||| | ||:: : || | ||||:||||
Qy   2739 GIEAPSVVRYTNAEILRVSQSLTDMTMPPLRAWRKKARAVLASAKRRGGAHAKLARFLL- 2797

Db   2888 GLSAFSLHSYSPGEINRVAACLRKLGVPPLRAWRHRARSVRARLLSRGGRAAICGKYLFN 2947
           | :|   |   |  |:||||||::| |||||||| ||:|||||| ||||| | |||||:
Qy   2798 WHATSR-----PLPDLDKTSVARYTTFNYCDVYSPEGDVFITPQRRLQKFLVKYLAVIVFA 2853

Db   2948 WAVRTKLKLTPIAAAGRLDLSGWFTAGYSG-----GDIYHSVSHARPRW------FWFC  2995
           | :: ||   ||  | |||| |||  ||      ||| |  || ||      ||||
Qy   2854 LGLIAVGLAI 2863
           | |:|| |:|
Db   2996 LLLLAAGVGI 3005
```

Search completed: October 27, 2005, 15:54:11
Job time : 323 secs

```
            GenCore version 5.1.6
      Copyright (c) 1993 - 2005  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    October 30, 2005, 08:53:15 ; Search time 2934 Seconds
                                         (without alignments)
                              18963.752 Million cell updates/sec

Title:          US-10-009-002-1
Perfect score:  9399
Sequence:       1 accacaaacactccagtttg.........cccgcttggaattaaaaact 9399

Scoring table:  IDENTITY NUC
                Gapop 10.0 , Gapext 1.0

Searched:       4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters:       8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_16Dec04:*
          1: geneseqn1980s:*
          2: geneseqn1990s:*
          3: geneseqn2000s:*
          4: geneseqn2001as:*
          5: geneseqn2001bs:*
          6: geneseqn2002as:*
          7: geneseqn2002bs:*
          8: geneseqn2003as:*
          9: geneseqn2003bs:*
         10: geneseqn2003cs:*
         11: geneseqn2003ds:*
         12: geneseqn2004as:*
         13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	9399	100.0	9399	9399	4	AAF23484	Aaf23484 GBV-B vir
2	9386.2	99.9	9399	9399	12	ADJ56732	Adj56732 Genome le
3	9386.2	99.9	9399	9399	12	ADJ64244	Adj64244 GB virus
4	9338.6	99.4	9397	9397		ADA77752	Ada77752 Genomic h
5	9131.4	97.2	9139	9139	4	AAF23485	Aaf23485 GBV-B vir
6	9059.8	96.4	9143	9143	2	AAT00130	Aat00130 Hepatitis
7	9059.8	96.4	9143	9143	3	AAA55376	Aaa55376 Hepatitis
8	9059.8	96.4	9143	9143	3	AAA55379	Aaa55379 Hepatitis
9	9037.4	96.2	9143	9143	2	AAT59785	Aat59785 Hepatitis
10	8692.6	92.5	8912	8912	2	AAA55280	Aaa55280 Hepatitis
11	8691.4	92.5	8912	8912	2	AAT00040	Aat00040 Hepatitis
12	6088	64.8	8069	8069	9	ADA77751	Ada77751 Hepatitis
13	4159	44.2	4268	4268	9	AAT00052	Aat00052 Hepatitis
14	4159	44.2	4268	4268	3	AAA55298	Aaa55298 Hepatitis
15	449.8	4.8	479	479	3	AAT00045	Aat00045 Hepatitis
16	449.8	4.8	479	479	3	AAA55378	Aaa55378 Hepatitis
17	449.8	4.8	479	479	2	AAA55291	Aaa55291 Hepatitis
c 18	319.6	3.4	337	337	2	AAT00127	Aat00127 Hepatitis
c 19	319.6	3.4	337	337	2	AAT00049	Aat00049 Hepatitis
c 20	319.6	3.4	337	337	3	AAA55371	Aaa55371 Hepatitis
c 21	319.6	3.4	337		3	AAA55295	AAA55295 Hepatitis
22	307.8	3.3	362		5	AAS15946	AAS15946 3' nontra
23	307.4	3.3	309		4	AAC92010	AAC92010 GBV-B 3'X
24	279.4	3.0	281		3	AAS55285	AAS55285 Hepatitis
25	279.4	3.0	281		3	AAS55286	AAS55286 Hepatitis
26	278.4	3.0	9405		3	AAQ40426	AAQ40426 Full-leng
27	273.6	2.9	9402		2	AAQ41345	AAQ41345 Human hep
c 28	273.2	2.9	9444		2	AAT13279	AAT13279 cDNA to g
29	270.4	2.9	9711		4	AAF23486	AAF23486 Infectiou
30	270.4	2.9	9711		4	AAC86937	AAC86937 Nucleotid
31	270.4	2.9	9711		5	AAC86644	AAC86644 Nucleotid
32	269.6	2.9	3564		2	AAQ32442	AAQ32442 HCV NS2-N
33	265.6	2.8	9589		2	AAQ38218	AAQ38218 NANBH vir
34	262.4	2.8	3970		2	AAQ38219	AAQ38219 NANBH vir
35	262.4	2.8	5211		4	AAC83408	AAC83408 DNA encod
36	260.2	2.8	9502		2	AAQ74770	AAQ74770 Hepatitis
37	257.4	2.7	259		4	AAC91998	AAC91998 GBV-B 3'X
38	256	2.7	8385		13	ADQ26514	ADQ26514 Hepatotro
39	256	2.7	9609		6	AAD33038	AAD33038 HCV-S1 fu
40	255.6	2.7	9436		2	AAQ63499	AAQ63499 Blood tra
41	254.4	2.7	260		12	ADJ56731	ADJ56731 3' termin
42	254.4	2.7	260		12	ADJ64243	ADJ64243 GB virus
43	253.4	2.7	7911		2	AAQ32436	AAQ32436 GB virus
44	252.8	2.7	3564		2	AAQ32501	AAQ32501 HCV NS2-N
45	252.6	2.7	2061		6	AAD34500	AAD34500 Hepatitis

ALIGNMENTS

```
RESULT 1
AAF23484
ID   AAF23484 standard; DNA; 9399 BP.
XX
AC   AAF23484;
XX
DT   11-SEP-2003  (revised)
DT   21-MAR-2001  (first entry)
XX
DE   GBV-B virus genome.
XX
KW   GBV-B; hepatitis C virus; HCV; vaccine; ds.
XX
OS   Hepatitis GB virus B.
XX
PN   WO200075337-A1.
XX
PD   14-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US015293.
XX
PR   04-JUN-1999;  99US-0137694P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Bukh J, Yanagi M, Emerson SU, Purcell RH;
XX
DR   WPI; 2001-091214/10.
XX
PT   New infectious nucleic acids of the GB virus-B clone, useful for
PT   indirectly studying the molecular properties of hepatitis C virus (HCV)
PT   and in developing vaccines and therapeutics for HCV.
XX
PS   Claim 3; Page 60-63; 96pp; English.
XX
CC   The present invention relates to GB virus-B. The nucleic acid molecules
CC   of the invention are useful for indirectly studying the molecular
CC   properties of hepatitis C virus (HCV). The infectious nucleic acid
CC   sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used
CC   in the development of vaccines and therapeutics for HCV. (Updated on 11-
CC   SEP-2003 to standardise OS field)
XX
SQ   Sequence 9399 BP; 2125 A; 2326 C; 2430 G; 2518 T; 0 U; 0 Other;
```

```
Query Match      100.0%;  Score 9399;  DB 4;  Length 9399;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 9399;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 ACCACAAACACTCCAGTTGTTACACTTCCGTCCCGCTAGGAATGCTCTGGAGCACCCCCTAG 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACCACAAACACTCCAGTTGTTACACTTCCGTCCCGCTAGGAATGCTCTGGAGCACCCCCTAG 60

Qy     61 CAGGGCTGGGGGATTTCCCCTGCCCGTCTGCAGAAGGGTGGAGCCAACCACCTTAGTAT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAGGGCTGGGGGATTTCCCCTGCCCGTCTGCAGAAGGGTGGAGCCAACCACCTTAGTAT 120

Qy    121 GTAGGCGCGGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTAGGCGCGGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC 180

Qy    181 CCTGATGGGCGTTCATGGGTTCGGTGGTGTGGTGGCGCGTTTAGGCAGCCTCCAGCCCACCA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCTGATGGGCGTTCATGGGTTCGGTGGTGTGGTGGCGCGTTTAGGCAGCCTCCAGCCCACCA 240

Qy    241 CCTCCCAGATAGAGCGCGGGCACTGTAGGGAAGACACCGGGGACCCGGTCACTACCAAGGACG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTCCCAGATAGAGCGCGGGCACTGTAGGGAAGACACCGGGGACCCGGTCACTACCAAGGACG 300

Qy    301 CAGACCTCTTTTGAGTATCACGCCTCCGGAGTAGTTGGGCAAGCCCACCTATATGTGT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGACCTCTTTTGAGTATCACGCCTCCGGAGTAGTTGGGCAAGCCCACCTATATGTGT 360

Qy    361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACCGCTGATAGGGGATCCTTGCCGAGGGGGAT 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACCGCTGATAGGGGATCCTTGCCGAGGGGGAT 420

Qy    421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTCTACTCAAACAAGTCCTGTACC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTCTACTCAAACAAGTCCTGTACC 480

Qy    481 TGCGCCCAGAAGACGCGCAAGAACAAGCAGACGCAGGCTTCATATCCTGTCATTAAAAC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGCGCCCAGAAGACGCGCAAGAACAAGCAGACGCAGGCTTCATATCCTGTCATTAAAAC 540

Qy    541 ATCTGTTGAAAGGGACAAAGTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ATCTGTTGAAAGGGACAAAGTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCG 600

Qy    601 TTACAAAATTGCTGGTATCGATGATGGCTTGCAGATAGTACGCTTGCAGTAGTACGCTTGCCAGC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTACAAAATTGCTGGTATCGATGATGGCTTGCAGATAGTACGCTTGCAGTAGTACGCTTGCCAGC 660

Qy    661 TCATGGTTGGGGACGGCCAAGACCCTCGGCCATAAGTCTCGCAATCTTCTGGAATCCTTCTGA 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TCATGGTTGGGGACGGCCAAGACCCTCGGCCATAAGTCTCGCAATCTTCTGGAATCCTTCTGA 720

Qy    721 TTACCCTTTGGGGTGGATTGGTGATGTTACAACTCACACCTCTAGGCCCGCTGGT 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TTACCCTTTGGGGTGGATTGGTGATGTTACAACTCACACCTCTAGGCCCGCTGGT 780

Qy    781 GGCAGGAGGCGGTCGTTCGACCAGTCTGCCAGATAGTACGCTTGCTGGAGGATGGAGTCAA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GGCAGGAGGCGGTCGTTCGACCAGTCTGCCAGATAGTACGCTTGCTGGAGGATGGAGTCAA 840

Qy    841 CTTGGGCTACTGGTTGGGTCGGTGTCCACCTTTTGTGTATGTCTGCTATCTTTGGCCTG 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTTGGGCTACTGGTTGGGTCGGTGTCCACCTTTTGTGTATGTCTGCTATCTTTGGCCTG 900

Qy    901 TCCCTGTAGTGGGGCGCGGTGGTTATCTATTGTTCTCCTTCCACTTGCCCACACCAATCCTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCCCTGTAGTGGGGCGCGGTGGTTATCTATTGTTCTCCTTCCACTTGCCCACACCAATCCTG 960

Qy    961 CTGCCAGCCAATCAGGTTATCATTGTTTCTCCTTCCACTTGCCTACACGAGCCTGGTTG 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CTGCCAGCCAATCAGGTTATCATTGTTTCTCCTTCCACTTGCCTACACGAGCCTGGTTG 1020

Qy   1021 TGTGATCTGTGCGGACGAGTGCTGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TGTGATCTGTGCGGACGAGTGCTGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA 1080

Qy   1081 TTGGACTGGCACGGACTCCTTCTTGGCTGACCACCATTGATTTTGTTATGGGCGCTTTGT 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TTGGACTGGCACGGACTCCTTCTTGGCTGACCACCATTGATTTTGTTATGGGCGCTCTTGT 1140

Qy   1141 GACCTGTGACGCCCTTGACATTGTGAGTGTGGTGCGTGTATTAGTCGGTGACTG 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTGTGACGCCCTTGACATTGTGAGTGTGGTGCGTGTATTAGTCGGTGACTG 1200

Qy   1201 GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGA 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGA 1260

Qy   1261 AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT 1320

Qy   1321 CGAGGCGTCATCTTCTTGACCAAAACTGGCTTCACAAGTACCATTAGCGGACTAT 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CGAGGCGTCATCTTCTTGACCAAAACTGGCTTCACAAGTACCATTAGCGGACTAT 1380

Qy   1381 GTTTAGCAGTGTACACTACCTGGCGCTTGGCGCTCTGATCTACTACATGCCTCTCGGGGCAA 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GTTTAGCAGTGTACACTACCTGGCGCTTGGCGCTCTGATCTACTACATGCCTCTCGGGGCAA 1440

Qy   1441 GTGGTATCAGTTGCTCCTAGCGCTTTATGCTTTACATAGAAGGCGACCTCTGGAAACCCCAT 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GTGGTATCAGTTGCTCCTAGCGCTTTATGCTTTACATAGAAGGCGACCTCTGGAAACCCCAT 1500

Qy   1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTTGCTCGCGCCTTTGATGATACCATGTCC 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTTGCTCGCGCCTTTGATGATACCATGTCC 1560

Qy   1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAGTGGAC 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAGTGGAC 1620

Qy   1621 CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATCAATCCCTG 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATCAATCCCTG 1680

Qy   1681 TGGCAGGGGATGTATGGTTAAAATTCAAAATAACACCATGGGTTGCTGCCGTATTCGCAA 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGGCAGGGGATGTATGGTTAAAATTCAAAATAACACCATGGGTTGCTGCCGTATTCGCAA 1740

Qy   1741 TGTGCCATCGTACTGCACTATGGCCACTGATGTCAGTGTGGAACGACACTCGCAACACTTA 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TGTGCCATCGTACTGCACTATGGCCACTGATGTCAGTGTGGAACGACACTCGCAACACTTA 1800

Qy   1801 CGAAGCATGCGGTGTAACACCATGGCTAACACCGGCTCAGCGGCTCAGCGCTGT 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CGAAGCATGCGGTGTAACACCATGGCTAACACCGGCTCAGCGGCTCAGCGCTGT 1860

Qy   1861 ATTGGCTATATTACAATACCCTGGGTCTCAAAGAAAATGTTTAAACCTTCATATATTGATGTC 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 ATTGGCTATATTACAATACCCTGGGTCTCAAAGAAAATGTTTAAACCTTCATATATTGATGTC 1920

Qy   1921 AGGCCATTTGTATTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AGGCCATTTGTATTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA 1980

Qy   1981 TTCACTCTTCCTACCACCGGAGAGGTGGCTACCCCACCTGTGTACG 2040
            ||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TTCCACTCTTCCTACCACCGGAGAGGTGGCTACCCCACCTGTGTACG 2040

Qy   2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGGTTTACAGTGATGTGAAAGAGACTAGCCACAGG 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGGTTTACAGTGATGTGAAAGAGACTAGCCACAGG 2100

Db   2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCCGCCACGGGTGC 2160
```

Left column:

```
Db  2101  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2101  ATTGATCACCAAGACAAAGCCTGAGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC  2160

Db  2161  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2161  TTTGTCTCTTACGGGAGTTACCACCACAGGCCGTGGTGCTAATTCTGTTGGGTTGTGTGG  2220

Db  2221  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2221  CAGCAGAAGTATCTTATTTAGGCTACCTCTGTTACTTGTTTGGGCGGCGCTTC  2280

Db  2281  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2281  TGGTTACCCTTTGCGTCTGTATCTCCAGTCGTATCTCCAAGCTGGCTGGGATGT  2340

Db  2341  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTTGCTTTGATTTTCTTCATCTGTTGCTATCTCCG  2400

Db  2401  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2401  CTGCAGGCTACGTTATGCTGCGCCAACCAGATATGACGTGGGTGCGACTGCT  2460

Db  2461  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2461  AACTTTCTTTGTTGCAGCAGTCTGCTGGCGCGGCCGCGTAACCGTGGGTGCGACTGCT  2520

Db  2521  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2521  AGTGGCAGGGTTAGTTTTGTTGTGGGGCCGGCCGCGTAACCGTGGTCACCCGGTAGCTCTGCTGCT  2580

Db  2581  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2581  AGGTCCTTGGCCTCTCGGTAGGGCCTTTAACCCCTCTGCATTGGTTACGGCCTTGCTTCAGC  2640

Db  2641  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2641  TTTTGATACCGAGATAATTGAGAGGGCTGACATAACCACCTGTAGTAGCATTGTCAT  2700

Db  2701  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2701  GTCTCGTTTTGGCTTTGCTCACTGTTACCTCGCTGCTTAGTTAACTCCTATCT  2760

Db  2761  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2761  TTGGCAACGTTGGGAGAATTGGTTTGGGAACGTTACACTAAGACCGGAGAGGTTTTCCT  2820

Db  2821  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2821  TGTGCTGGTTGTTCCCCGTGGCGACATATGCAGCGCCTGTGACTTTCTGTGTGTCA  2880

Db  2881  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2881  CGTAGCTCTTCTATGTTAACATCCAGTGCAGCATCGTTCTTTGGGACTGACTCTAGGGT  2940

Db  2941  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2941  TAGGGCCCATAGAATGTTGGTGCGTCTCGGAAGTGTCATGCTTGGTATTCTCATTATGT  3000

Db  3001  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3001  TCTTAAGTTTTTCCTCTTAGTGTTTTGGTTGGTGAGAATGGTGTGTTTTTCTATAAGCACTTGCA  3060

Db  3061  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3061  TGGTGATGGTCTTGCCTAATGAGTTTGCCTCGAAACTACCATTGCAAGAGCCATTTTCCC  3120

Db  3121  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGAAGGTCAAGGAATGAAGGAAGAAGACGCTTGGCGTGTGGGGACAC  3180

Db  3181  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3181  GGTTGATGGTTGCCCGTTGTGCGGTCTCGGCGGACCTTGTTTCGCAGGGTTGGCTAT  3240
```

Right column:

```
Db  3181  GGTTGATGGTTGCCCGTTGTGCGGTCTCGGCGGACCTTGTTTCGCAGGGTTGGCTAT  3240
Db  3241  GCCGCCAGATGGGTGGGCCATTACCACCACGACCTTTACGCTGCAGTGTCTCTGAACGTGG  3300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3241  GCCGCCAGATGGGTGGGCCATTACCACCACGACCTTTACGCTGCAGTGTCTCTGAACGTGG  3300

Db  3301  CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCCGAACTGGACTGGAAC  3360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3301  CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCCGAACTGGACTGGAAC  3360

Db  3361  TATCTTCAGATTAGGATCTCTGCCACTAGCTACACATGGGATTTGTTGTGACAACGTGTT  3420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3361  TATCTTCAGATTAGGATCTCTGCCACTAGCTACACATGGGATTTGTTGTGACAACGTGTT  3420

Db  3421  GTATACTGCTCACCATGCGCAGCGCCGTTGGCTCATCCCACCACCATGTCTATACA  3480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3421  GTATACTGCTCACCATGCGCAGCGCCGTTGGCTCATCCCACCACCATGTCTATACA  3480

Db  3481  CCCAATAACCGTTGACGGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG  3540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3481  CCCAATAACCGTTGACGGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG  3540

Db  3541  GTCCCTACTCGGTGCTCTTGCGGTGTGACCAAGGGTATCGTGCGGGGCCCTTCCCAT  3600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3541  GTCCCTACTCGGTGCTCTTGCGGTGTGACCAAGGGTATCGTGCGGGGCCCTTCCCAT  3600

Db  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCCTATTGGTGTGTGCGGGGCCCTTCCCAT  3660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCCTATTGGTGTGTGCGGGGCCCTTCCCAT  3660

Db  3661  GGCTGTTGCCAAGGGTTCTTCAGGTGCCCGGATTCTGTGCTCTCCGGCATGTTATTGG  3720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3661  GGCTGTTGCCAAGGGTTCTTCAGGTGCCCGGATTCTGTGCTCTCCGGCATGTTATTGG  3720

Db  3721  GATGTCACCGGCTGCTAGAAATTCTGGCCGGTTCAGTCAGTCAGATTAGGGTTAGGCCGTT  3780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3721  GATGTCACCGGCTGCTAGAAATTCTGGCCGGTTCAGTCAGTCAGATTAGGGTTAGGCCGTT  3780

Db  3781  GGTGTGTCTGGATACCATCCCGCTGCTTAGTTAGTTCACTCTTGATACAAAACCTAC  3840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3781  GGTGTGTCTGGATACCATCCCGCTGCTTAGTTAGTTCACTCTTGATACAAAACCTAC  3840

Db  3841  TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC  3900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3841  TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC  3900

Db  3901  CAAATTACCACTTCTTCTTACATGCAGAGGAAGTGATGAGTCTTGGCTCCTAAATCCCAGTGT  3960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3901  CAAATTACCACTTCTTCTTACATGCAGAGGAAGTGATGAGTCTTGGCTCCTAAATCCCAGTGT  3960

Db  3961  GGCTACAACAGCATCAATGCCAAAGTACATGACGCGACGTACGGGGTGAATCCAAATTG  4020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3961  GGCTACAACAGCATCAATGCCAAAGTACATGACGCGACGTACGGGGTGAATCCAAATTG  4020

Db  4021  CTATTTTAATGCAAATGTACAACAACAGGGGCTTCACTTACGTACAGCACATATGGCAT  4080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4021  CTATTTTAATGCAAATGTACAACAACAGGGGCTTCACTTACGTACAGCACATATGGCAT  4080

Db  4081  GTACCTGACCGGACGCATGTTCCCGGTTGGCATTGGGATGTAATCATTGTGACGAATGCCATGC  4140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4081  GTACCTGACCGGACGCATGTTCCCGGTTGGCATTGGGATGTAATCATTGTGACGAATGCCATGC  4140

Db  4141  TACCGATGCAACCACCGTGTTGCCTCGAAACTACCATTGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA  4200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4141  TACCGATGCAACCACCGTGTTGCCTCGAAACTACCATTGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA  4200

Db  4201  TGTTAGGCTAGTGGTTCTTGCCACGGCTTCGCCGTGTGGGGACAC  4260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4201  TGTTAGGCTAGTGGTTCTTGCCACGGCTTCGCCGTGTGGGGACAC  4260

Db  4261  CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAGAT  4320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4261  CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAGAT  4320
```

```
Qy  4321  TAAGGAGAAATCTGAAGAAGGGAGACACCTTATCTTTGAGGCTACCAAAAAAACACTG  4380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321  TAAGGAGAAATCTGAAGAAGGGAGACACCTTATCTTTGAGGCTACCAAAAAAACACTG  4380

Qy  4381  TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG  4440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381  TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG  4440

Qy  4441  ATGTGACATCTCAAAAATCCCTGAGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG  4500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441  ATGTGACATCTCAAAAATCCCTGAGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG  4500

Qy  4501  TACAGGGTACACTGGTGACTTGATTCCGTGTATGACTGCAGCCTCATGGTAGAAGGCAC  4560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501  TACAGGGTACACTGGTGACTTGATTCCGTGTATGACTGCAGCCTCATGGTAGAAGGCAC  4560

Qy  4561  ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTGTTCGTGTGTGCGGGGTTTCAGC  4620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561  ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTGTTCGTGTGTGCGGGGTTTCAGC  4620

Qy  4621  AATAGTTAAAGGCCAGCGTAGGGCCGTTGGGGAGAGCTGGCATATACTACTA  4680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621  AATAGTTAAAGGCCAGCGTAGGGCCGTTGGGGAGAGCTGGCATATACTACTA  4680

Qy  4681  TGTGAGACGGGAGTTGTACCCCCTTCGGGTATGTTCCTGAATGCAACATTGTTGAAGCTT  4740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681  TGTGAGACGGGAGTTGTACCCCCTTCGGGTATGTTCCTGAATGCAACATTGTTGAAGCTT  4740

Qy  4741  CGACGCAGCGCAAGGCATGTGTATGGTTGTCATCAACAACTACAAACTATTCTGGACAC  4800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741  CGACGCAGCGCAAGGCATGTGTATGGTTGTCATCAACAACTACAAACTATTCTGGACAC  4800

Qy  4801  CTATCGCACCCAACCTGGGTTACCTGCGGATAGGAGCAAATTTGACGAGTGGGCTGATCT  4860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801  CTATCGCACCCAACCTGGGTTACCTGCGGATAGGAGCAAATTTGACGAGTGGGCTGATCT  4860

Qy  4861  CTTTTCTATGGTCAACCCCGACCTTCATTGTCAATACTGCAAAAAGAACTGCTGACAA  4920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861  CTTTTCTATGGTCAACCCCGACCTTCATTGTCAATACTGCAAAAAGAACTGCTGACAA  4920

Qy  4921  TTATGTTTGTTGTACTGCAGCGCCAACTACAACTGTGTCATCAGTATGGCTATGCTCGCC  4980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921  TTATGTTTGTTGTACTGCAGCGCCAACTACAACTGTGTCATCAGTATGGCTATGCTCGCC  4980

Qy  4981  CAATGACGCACCACGGTCGGCAGGGAGCCCGGCTTGGGCCCGGCCGTTGGCGCGTTCTGTG  5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981  CAATGACGCACCACGGTCGGCAGGGAGCCCGGCTTGGGCCCGGCCGTTGGCGCGTTCTGTG  5040

Qy  5041  GCGCTTGACGGGCGCGACGGCTGACGTCCTGTCCTGGCCCCAGCGCCAGCGGGTGACCAGATACCA  5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041  GCGCTTGACGGGCGCGACGGCTGACGTCCTGTCCTGGCCCCAGCGCCAGCGGGTGACCAGATACCA  5100

Qy  5101  AATGTGCTTCACTGAAGTCAATACTTCTGGGACACTTCTGGACAGCGCACCTCGCTGTTGGCGTTGGGAGT  5160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101  AATGTGCTTCACTGAAGTCAATACTTCTGGGACACTTCTGGACAGCGCACCTCGCTGTTGGCGTTGGGAGT  5160

Qy  5161  GGCTATGGCTTATCTGACAACACTTTTGGCGCGCACTTTGGCGGCGCCACTTGTGGCGCGGTTGCTGTGTC  5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161  GGCTATGGCTTATCTGACAACACTTTTGGCGCGCACTTTGGCGGCGCCACTTGTGGCGCGGTTGCTGTGTC  5220

Qy  5221  TATTACATCAGTCCCTACCGGGTGCTACTGTCGCGCCCCAGTGGTTGACGAAGAAGAAATCGT  5280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221  TATTACATCAGTCCCTACCGGGTGCTACTGTCGCGCCCCAGTGGTTGACGAAGAAGAAATCGT  5280

Qy  5281  GGAGGAGTGCATCATTCATTCCCTGGAGGGCCATGGTTGCTGCAATTGACAGCTGAA  5340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281  GGAGGAGTGCATCATTCATTCCCTGGAGGGCCATGGTTGCTGCAATTGACAGCTGAA  5340

Qy  5341  GAGTACAATCACCACAACTAGTCCTTTCACATTGGAAACCGCCCTTGAAAAACTTAACAC  5400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341  GAGTACAATCACCACAACTAGTCCTCTCCAAATGTCTGTTTCACCCAGTGCCCCCAAC  5400
```

```
Qy  5401  CTTTCTTGGGCCTCATGCGAGCTACAAATCCTTGCTATCATAGAGTATTGCTGTGTTTAGT  5460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401  CTTTCTTGGGCCTCATGCGAGCTACAAATCCTTGCTATCATAGAGTATTGCTGTGTTTAGT  5460

Qy  5461  CACTTACCTGACAATCCCTTGCATCATGCGTGTTGCTTTCATTGCGGTATTACTAC  5520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461  CACTTTACCTGACAATCCCTTTGCATCATGCGTGTTTGCTTTCATTGCGGGATTACTAC  5520

Qy  5521  CCCACTACCTCACAGATCAAAAATGTCCGTGCATTATTGGAGGCGCAATTGGCGTCCAA  5580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521  CCCACTACCTCACAGATCAAAAATGTCCGTGCATTATTTGGAGGCGCAATTGGCGTCCAA  5580

Qy  5581  GCTTACAGACGCTAGAGCGCACTGGCGTTCATGATGCCGGGCTCGGGGACAGCTCT  5640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581  GCTTACAGACGCTAGAGCGCACTGGCGTTCATGATGCCGGGCTGCGGGACAGCTCT  5640

Qy  5641  TGGTACATGGACATCGGTGGTTTGTCTTGACATTTAAATGCTGATGGTGAGTGGCCACTATCAGCT  5700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641  TGGTACATGGACATCGGTGGTTTGTCTTGACATTTAAATGCTGATGGTGAGTGGCCACTATCAGCT  5700

Qy  5701  ATCCACTGCTTGCTTGACACTTCGGTGGTTTGTCTTGATATCCGCGGCTCAATCCGGGCGTCTTGTGTCGCAGC  5760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701  ATCCACTGCTTGCTTGACACTTCGGTGGTTTGTCTTGATATCCGCGGCTCAATCCGGGCGTCTTGTGTCGCAGC  5760

Qy  5761  TGCTGGTTTAGTCTACTCCGGGTTCAATCCGCGGCCGCAGGAGTTGTGGGCGTGTGGGCGTCTTGTGTCTTAC  5820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761  TGCTGGTTTAGTCTACTCCGGGTTCAATCCGCGGCCGCAGGAGTTGTGGGCGTGTGGGCGTCTTGTGTCTCAGC  5820

Qy  5821  TTGTGCAATGTTGCTTTGACAACAGCAGGGCCAGATCACTGGCCCAACAGACTTCTTAC  5880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821  TTGTGCAATGTTTGCTTTGTCTTTGACAACAGCAGGGCCAGATCACTGGCCCAACAGACTTCTTAC  5880

Qy  5881  TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCCACTCGTGACATCCG  5940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881  TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG  5940

Qy  5941  CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT  6000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941  CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT  6000

Qy  6001  CCGTTGGCTCCACACCCCGACGGAGGATGATTGATTGCGGGCTCATTGCTTTCATTGGGTCTCAGAGAT  6060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001  CCGTTGGCTCCACACCCCGACGGAGGATGATTGATTGCGGGCTCATTGCTTTCATTGGGTCTCAGAGAT  6060

Qy  6061  TTTGCAGTATGTGTGCAAATTCTTGTGTGATTTGTTAATGTCCTTAATGTCCTTAAAGCTGGAGTTCA  6120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061  TTTGGCAGTATGTGTGCAAATTTCTTGTGTGATTTGCTTTAATGTCCTTAAAGCTGGAGTTCA  6120

Qy  6121  GAGCATGGTTAACATTCCTGGTTGTCCTTTCACAGCTGCCAGAAGGGGTACAAGGGCCC  6180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121  GAGCATGGTTAACATTCCTGGTTGTCCTTTCTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC  6180

Qy  6181  CTGGATTGGATCAGGTATGCTCCAAGCACGCGTGTCCAAGCACGCGTCCATGCGGTGCTGAACTCATCTTTTC  6240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181  CTGGATTGGATCAGGTATGCTCCAAGCACGCGTGTCCAAGCACGCGTCCATGCGGTGCTGAACTCATCTTTTC  6240

Qy  6241  TGTTGAGAATGGTTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241  TGTTGAGAATGGTTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300

Qy  6301  AGGGGCTGTTCCAGTCAACGCTACGGCTGTGTGGGTCGGCTAGACCGGACCAACTGATTG  6360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301  AGGGGCTGTTCCAGTCAACGCTACGGCTGTGTGGGTCGGCTAGACCGGACCAACTGATTG  6360

Qy  6361  GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAATGGGAGA  6420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361  GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAATGGGAGA  6420

Qy  6421  TCACATTTTTGTTACAGCTAGTATCCTCTCCAAATGTCTGTTTCACCCAGTGCCCCCAAC  6480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421  TCACATTTTTGTTACAGCTAGTATCCTCTCCAAATGTCTGTTTCACCCAGTGCCCCCAAC  6480

Qy  6481  CTTGAGAGCTTGCAGTGGCCGTGACGGGCTACAGGTTCAGTGTTATCTCAGTGTTAGGTTGGC  6540
```

us-10-009-002-1.rng

Mon Oct 31 10:27:43 2005

```
Qy  8701  GATCCTCAGAGTTCCAATCACTAACAGACATGACCATGCCCCCCTGCGAGCCCTGCGCG  8760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8701  GATCCTCAGAGTTCCAATCACTAACAGACATGACCATGCCCCCCTGCGAGCCCTGCGCG  8760

Qy  8761  AAAGAAAGCCAGGGCGGTCCTGCCCAGAGGCGTGGCGGAGCACACGCAAAATT        8820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8761  AAAGAAAGCCAGGGCGGTCCTGCCCAGAGGCGTGGCGGAGCACACGCAAAATT        8820

Qy  8821  GGCTCGCTTCCTTCTCTGGCATGCTACATCTAGACCTTCTACCAGATTTGGATAAGACGAG  8880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8821  GGCTCGCTTCCTTCTCTGGCATGCTACATCTAGACCTTCTACCAGATTTGGATAAGACGAG  8880

Qy  8881  CGTGGCTCGGTACACCACTTCAATTATTGTGATGTTTACTCCCCGGAGGGGGATGTGTT   8940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8881  CGTGGCTCGGTACACCACTTCAATTATTGTGATGTTTACTCCCCGGAGGGGGATGTGTT   8940

Qy  8941  TATTACCACCAGAGAGAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTGTCATTGTTT  9000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8941  TATTACCACCAGAGAGAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTGTCATTGTTT  9000

Qy  9001  TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA   9060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9001  TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA   9060

Qy  9061  CTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGCAACAGGGGAGACCCCGGGC 9120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9061  CTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGCAACAGGGGAGACCCCGGGC 9120

Qy  9121  TTAACGACCCCCGCGATGTGAGTTGGGACCATGGTTGGCGACCATGGTTGGATCAGAAC   9180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9121  TTAACGACCCCCGCGATGTGAGTTGGGACCATGGTTGGCGACCATGGTTGGATCAGAAC   9180

Qy  9181  GCCATGGTCTGAAGGGGATGAGCGTCCCTTCTGGCTCATCCACAAAAAACCGTCTCGGGTGG  9240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9181  GCCATGGTCTGAAGGGGATGAGCGTCCCTTCTGGCTCATCCACAAAAAACCGTCTCGGGTGG  9240

Qy  9241  GTGAGAGTCCTGGCTGTGTGGAAGCAGTCAGTATAATTCCCGTGTGTGTGGTGACGC     9300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9241  GTGAGAGTCCTGGCTGTGTGGAAGCAGTCAGTATAATTCCCGTGTGTGTGGTGACGC     9300

Qy  9301  CTCACGACGTATTTGTCCGCTGCAGAGCGTAGTACCCAAGGGCTGCACCCCGGTTTTTG   9360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9301  CTCACGACGTATTTGTCCGCTGCAGAGCGTAGTACCCAAGGGCTGCACCCCGGTTTTTG   9360

Qy  9361  TTCCAAGCGGAGGGCAACCCCCGCTTGGAATTAAAAACT  9399
          |||||||||||||||||||||||||||||||||||||||
Db  9361  TTCCAAGCGGAGGGCAACCCCCGCTTGGAATTAAAAACT  9399
```

```
RESULT 2
ADJ56732
ID   ADJ56732 standard; cDNA; 9399 BP.
XX
AC   ADJ56732;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Genome length hepatitis GB virus B cDNA SeqID 2.
XX
KW   ss; GB virus B; GBV-B; HCV; flavivirus; hepatitis C virus; antiviral;
KW   vaccine; virucidal; antiinflammatory.
XX
OS   Hepatitis GB virus B.
XX
FH   Key            Location/Qualifiers
FH
FT   stem_loop      4..21
FT                  /*tag= a
FT   stem_loop      29..61
FT                  /*tag= b
FT   misc_binding   63..73
FT                  /*tag= c
FT                  /bound_moiety= "GBV-B DNA bases 236-226"
FT   stem_loop      81..131
FT                  /*tag= d
FT   stem_loop      135..154
FT                  /*tag= e
FT   stem_loop      156..218
FT                  /*tag= f
FT   misc_binding   226..236
FT                  /*tag= g
FT                  /bound_moiety= "GBV-B DNA bases 73-63"
FT   misc_binding   242..250
FT                  /*tag= h
FT                  /bound_moiety= "GBV-B DNA bases 427-419"
FT   misc_binding   251..257
FT                  /*tag= i
FT                  /bound_moiety= "GBV-B DNA bases 394-388"
FT   misc_binding   258..270
FT                  /*tag= j
FT                  /bound_moiety= "GBV-B DNA bases 362-349"
FT   stem_loop      273..287
FT                  /*tag= k
FT   stem_loop      288..337
FT                  /*tag= l
FT   stem_loop      339..348
FT                  /*tag= m
FT   misc_binding   349..362
FT                  /*tag= n
FT                  /bound_moiety= "GBV-B DNA bases 270-258"
FT   stem_loop      363..384
FT                  /*tag= o
FT   misc_binding   388..394
FT                  /*tag= p
FT                  /bound_moiety= "GBV-B DNA bases 257-251"
FT   stem_loop      395..406
FT                  /*tag= q
FT   misc_binding   410..415
FT                  /*tag= r
FT                  /bound_moiety= "GBV-B DNA bases 434-429"
FT   misc_binding   419..427
FT                  /*tag= s
FT                  /bound_moiety= "GBV-B DNA bases 250-242"
FT   misc_binding   429..434
FT                  /*tag= t
FT                  /bound_moiety= "GBV-B DNA bases 415-410"
FT   stem_loop      435..459
FT                  /*tag= u
XX
PN   WO2004005498-A1.
XX
PD   15-JAN-2004.
XX
PF   02-JUL-2003; 2003WO-US021002.
XX
PR   03-JUL-2002; 2002US-00189359.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
PA   (INSP ) INST PASTEUR.
XX
PI   Martin A, Sangar DV, Lemon SM, Rijnbrand R;
XX
DR   WPI; 2004-091362/09.
XX
PT   New chimeric GBV-B polynucleotide, useful as a model for hepatitis C
PT   virus, for identifying compounds active against a viral infection, or for
PT   developing hepatitis C virus preventive and therapeutic treatments.
XX
PS   Example 22; SEQ ID NO 2; 108pp; English.
XX
CC   This invention relates to novel isolated chimeric GB virus B (GBV-B)/HCV
CC   polynucleotides. Specifically, it refers to using the hepatotropic
CC   flavivirus GBV-B that has a unique phylogenetic relationship to the human
CC   hepatitis C virus (HCV) and can serve as a surrogate virus in drug
CC   discovery efforts related to antiviral drug development. The present
CC   invention describes the construction of an infectious molecular clone
```

```
CC   using the newly determined 3' terminal sequence of GBV-B. Furthermore,
CC   the GBV-B/HCV chimeras exhibit liver-specific expression and express HCV
CC   envelope proteins such that they can have utility as a vaccine immunogen
CC   for hepatitis C. In addition, they can be used for screening compounds
CC   active against viral infection, as well as for developing HCV
CC   preventative and therapeutic treatments. Accordingly, these compositions
CC   exhibit virucidal, antiinflammatory and hepatotropic activities. This
CC   polynucleotide sequence is the GBV-B cDNA sequence of the invention.
XX
SQ   Sequence 9399 BP; 2124 A; 2322 C; 2432 G; 2521 T; 0 U; 0 Other;

Query Match      99.9%;  Score 9386.2;  DB 12;  Length 9399;
Best Local Similarity  99.9%;  Pred. No. 0;
Matches 9391;  Conservative   0;  Mismatches   0;  Indels   8;  Gaps   0;

Qy      1  ACCAAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG    60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  ACCAAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG    60

Qy     61  CAGGGGCGTGGGGGATTTCCCCGCCCGTCTGCGAGAAGGGTGAGCCAACCCACCTTAGTAT   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CAGGGGCGTGGGGGATTTCCCCGCCCGTCTGCGAGAAGGGTGAGCCAACCCACCTTAGTAT   120

Qy    121  GTAGGCGCGGGGGACTCATGACGCTCGGCTGATGACAGCGCCAAGCTTGACTTGGATGGC   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GTAGGCGCGGGGGACTCATGACGCTCGGCTGATGACAGCGCCAAGCTTGACTTGGATGGC   180

Qy    181  CCTGATGGCCGTTCATGGGTTCGGTGGTGGTGGCGCCTTAGGCAGCCTCCACGCCCACCA   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  CCTGATGGGCCGTTCATGGGTTCGGTGGTGGTGGCGCCTTAGGCAGCCTCCACGCCCACCA   240

Qy    241  CCTCCCAGATAGAGCCGGCGGCCACTGTAGGGGAAGACCGGGTCACTACCAAGGACG   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  CCTCCCAGATAGAGCCGGCGGCCACTGTAGGGGAAGACCGGGTCACTACCAAGGACG   300

Qy    301  CAGACCTCTTTTGAGTATCACGCCTCCGGAGTAGTGGGCAAGCCCACCTATATGTGT   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  CAGACCTCTTTTGAGTATCACGCCTCCGGAGTAGTGGGCAAGCCCACCTATATGTGT   360

Qy    361  TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCTGATAGGGGTCCTTGCGAGGGGAT   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCTGATAGGGGTCCTTGCGAGGGGAT   420

Qy    421  CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTATTTCTACTCAAACAAGTCCTGTACC   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTATTTCTACTCAAACAAGTCCTGTACC   480

Qy    481  TGCGCCCAGAACGCGCAAGAACAAGCAGACGCAGGCTTCATATCCTGTGTCATTAAAAC   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  TGCGCCCAGAACGCGCAAGAACAAGCAGACGCAGGCTTCATATCCTGTGTCATTAAAAC   540

Qy    541  ATCTGTTGAAAGGGGACAACGAGCAAAGCCAAAGTCAGCGCGATGCTCCGGCCTCGTAA   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  ATCTGTTGAAAGGGGACAACGAGCAAAGCCAAAGTCAGCGCGATGCTCCGGCCTCGTAA   600

Qy    601  TTACAAAATTGCTGGTATCCATGATGCTTCAGGCTGCTTTGCCAGC   660
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    601  TTACAAAATTGCTGGTATCCATGATGCTTCAGGCTGCTTTGCCAGC   660

Qy    661  TCATGGTTGGGGACGCGTCGTTGCCCAATCCTGGAATCCTTCTGGA   720
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    661  TCATGGTTGGGGACGCGTCGTTGCCCAATCCTGGAATCCTTCTGGA   720

Qy    721  TTACCCTTTGGGGTGGACGCGTCGTTGCCGACCAGTCTGCCAGATAGTACGCTTCGGAGGATGGAGTCAA   780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  TTACCCTTTGGGGTGGACGCGTCGTTGCCGACCAGTCTGCCAGATAGTACGCTTCGGAGGATGGAGTCAA   780

Qy    781  GGCAGGAGCGGTCGTTGCCGACCAGTCTCGCCAGATACGCGTTGCTGCTCGGAGGATGGAGTCAA   840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  GGCAGGAGCGGTCGTTGCCGACCAGTCTCGCCAGATACGCGTTGCTGCTCGGAGGATGGAGTCAA   840

Qy    841  CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTGTGGTATGTCTGCTATCTTGGCCTG   900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTGTGGTATGTCTGCTATCTTGGCCTG   900

Qy    901  TCCCTGTAGTGGGGCGGCGGGGTCACTGACCAGACAAATACACCAATCCTGACCAATTG   960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  TCCCTGTAGTGGGGCGGCGGGGTCACTGACCAGACAAATACACCAATCCTGACCAATTG   960

Qy    961  CTTGCCAGCGTAATCAGTTATCTATTGTTCTCCTTCCACTTGCCTCACGAGCCTGGTTG  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  CTTGCCAGCGTAATCAGTTATCTATTGTTCTCCTTCCACTTGCCTCACGAGCCTGGTTG  1020

Qy   1021  TGTGATCGTGTGCGGACGAGTGGTGGGTTCCCGGTCCCCGCCAATCCGTACACCTTCCAA  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  TGTGATCGTGTGCGGACGAGTGGTGGGTTCCCGGTCCCCGCCAATCCGTACACCTTCCAA  1080

Qy   1081  TTGGACTGGCCACGGACTCCTTCTGCTGCGTGACCACATTGATTTTGTTATGGGCGCTCTTGT  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  TTGGACTGGCCACGGACTCCTTCTGCTGCGTGACCACATTGATTTTGTTATGGGCGCTCTTGT  1140

Qy   1141  GACCTGTGACGCCCTTGACATTGGTGAGTGTGTGCGTGTATTAGTCGGTGACTG  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  GACCTGTGACGCCCTTGACATTGGTGAGTGTGTGCGTGTATTAGTCGGTGACTG  1200

Qy   1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTACCTGA  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTACCTGA  1260

Qy   1261  AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT  1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT  1320

Qy   1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT  1380

Qy   1381  GTTTAGCAGTGTACACTACCTGGCGGTTGGCGCTCGATCTGATCTACTACTATATGTGT  1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381  GTTTAGCAGTGTACACTACCTGGCGGTTGGCGCTCGATCTGATCTACTACTATATGTGT  1440

Qy   1441  GTGGTATCAGTTGCTCCTCAGCGCTTATGCTTACATAGAAGCGACCTCTGGAAACCCCAT  1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441  GTGGTATCAGTTGCTCCTCAGCGCTTATGCTTACATAGAAGCGACCTCTGGAAACCCTAT  1500

Qy   1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCTTGATGATACCATGTCC  1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCTTGATGATACCATGTCC  1560

Qy   1561  TTGCCACTCTTATTGAGTGAGAATGTGTCAGAGTCATTGTTACAGTCCAAAGTGGAC  1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561  TTGCCACTCTTATTGAGTGAGAATGTGTCAGAGTCATTGTTACAGTCCAAAGTGGAC  1620

Qy   1621  CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG  1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621  CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG  1680

Qy   1681  TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGTGCTGCCGTATTCGCAA  1740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681  TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGTGCTGCCGTATTCGCAA  1740

Qy   1741  TGTGCCATCGTACTGCACTGATCGATGCAGTGTGGAACGACACTCGCAACACTTA  1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741  TGTGCCATCGTACTGCACTGATCGATGCAGTGTGGAACGACACTCGCAACACTTA  1800

Qy   1801  CGAAGCATGCGGTGTAACACGTGTCAACCGCTAACACCGCTCAGCCCTGA  1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801  CGAAGTATGCGGTGTAACACATGGTCAACACCGCTAACACCGCTCAGCCCTGA  1860

Qy   1861  ATTGGCTATATTACAATCACCCTGGTCTAAAGAAAATGTTTAAACCTCATAATTGGATGTC  1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861  ATTGGCTATATTACAATCACCCTGGTCTAAAGAAAATGTTTAAACCTCATAATTGGATGTC  1920

Qy   1921  AGGCCATTTGTATTTTGAGGGATCAGATACCCCTATAGTTTACTTTTACTTTTATGACCCTGTGAA  1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921  AGGCCATTTGTATTTTGAGGGATCAGATACCCCTATAGTTTACTTTTACTTTTATGACCCTGTGAA  1980
```

```
Qy 1981 TTCCACTCTCCTACCACGGAGAGGTGGGCTAGGTTGCCCGGTACCCACCTGTGGTACG 2040
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981 TTCCACTCTCCTACCACGGAGAGGTGGGCTAGGTTGCCCGGTACCCACCTGTGGTACG 2040

Qy 2041 TGGTTCTTGGTTACAGGTTCCCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041 TGGTTCTTGGTTACAGGTTCCCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100

Qy 2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC 2160
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC 2160

Qy 2161 TTTGTCTCTTACGGGAGTTACCACCAAGGCCGGTGTGCTAATTCTGTTGGGGTTGTGTGG 2220
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161 TTTGTCTCTTACGGGAGTTACCACCAAGGCCGGTGTGCTAATTCTGTTGGGGTTGTGTGG 2220

Qy 2221 CAGCAAGTATCTTATTTTAGCTGTACCTTGTTCCCTTTGTTTTGGGCGCGCCTTC 2280
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221 CAGCAAGTATCTTATTTTAGCTGTACCTTGTTCCCTTTGTTTTGGGCGCGCCTTC 2280

Qy 2281 TGGTTACCCTTGCGTCCTGTCTCCATCCCAGTGTATCTTCCAAGTGGCTGGGATGT 2340
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2281 TGGTTACCCTTGCGTCCTGTCTCCATCCCAGTGTATCTTCCAAGTGGCTGGGATGT 2340

Qy 2341 TTTGTCTAAAGCTCAAGTAGCTCCTTTGCTTTGATTTTCTTCATCTGTTGTATCTCCG 2400
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2341 TTTGTCTAAAGCTCAAGTAGCTCCTTTGCTTTGATTTTCTTCATCTGTTGTATCTCCG 2400

Qy 2401 CTGCAGGCTACGTTATGCGTGCCCTTTAGGGTTTGTGCCCATGGCGGCCTGCGCCCCT 2460
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401 CTGCAGGCTACGTTATGCGTGCCCTTTAGGGTTTGTGCCCATGGCGGCCTGCGCCCCT 2460

Qy 2461 AACTTTCTTGTTGCAGCAGCTCGTGCCGCCGCTGCTCCTTGCATTGGTGTTGCGACTGCT 2520
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2461 AACTTTCTTGTTGCAGCAGCTCGTGCCGCCGCTGCTCCTTGCATTGGTGTTGCGACTGCT 2520

Qy 2521 AGTGGCAGGGTAGTTTGTTGTGGGGCCGGCCGGTGTAACCGTGGTCACCGCATAGCTCTGCTGT 2580
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2521 AGTGGCAGGGTAGTTTGTTGTGGGGCCGGCCGGTGTAACCGTGGTCACCGCATAGCTCTGCTGT 2580

Qy 2581 AGGTCCTTGGCCTCTGGAGCGCGTTTTAACCCTCCTTGCATTGGTTACGCGCTTCTTCAGC 2640
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2581 AGGTCCTTGGCCTCTGGAGCGCGTTTTAACCCTCCTTGCATTGGTTACGCGCTTCTTCAGC 2640

Qy 2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT 2700
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT 2700

Qy 2701 GTCTCGTTTTGGCTTCTTGCTTCTTTGCTCACTTGTTACCTCGCTGTGTCTTAGTTAACTCCTATCT 2760
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2701 GTCTCGTTTTGGCTTCTTGCTTCTTTGCTCACTTGTTACCTCGCTGTGTCTTAGTTAACTCCTATCT 2760

Qy 2761 TTTGGCAACGTTGGGAGAATTGGTTTGGAACCGTTACACTAAGACCGGAGAGGGTTTTCCT 2820
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761 TTTGGCAACGTTGGGAGAATTGGTTTGGAACCGTTACACTAAGACCGGAGAGGGTTTTCCT 2820

Qy 2821 TGTGCTGTTGTTGTTTTCCCCGGTGGACATATGACGCGCTGGTGACTTTCTGTGTGTGTCA 2880
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2821 TGTGCTGTTGTTGTTTTCCCCGGTGGACATATGACGCGCTGGTGACTTTCTGTGTGTGTCA 2880

Qy 2881 CGTAGCTCTTCTTGGCTTCTATGTTTAACATCCAGTGCAGCATCGTTCTTTTGGGACTGACTCTAGGGT 2940
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2881 CGTAGCTCTTCTTGGCTTCTATGTTTAACATCCAGTGCAGCATCGTTCTTTTGGGACTGACTCTAGGGT 2940

Qy 2941 TAGGGCCATAGAAATGGTTGGTCTCGGAAAGTGTCATGCTTGGTATTCTTCATTATGT 3000
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2941 TAGGGCCATAGAAATGGTTGGTCTCGGAAAGTGTCATGCTTGGTATTCTTCATTATGT 3000

Qy 3001 TCTTAAGTTTTTCCCTCTTAGTGTTTGGTGAGAATGGTGTGTTTCTATAAGACTTGCA 3060
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3001 TCTTAAGTTTTTCCCTCTTAGTGTTTGGTGAGAATGGTGTGTTTCTATAAGACTTGCA 3060

Qy 3061 TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCAAGAGCCATTTTCCC 3120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3061 TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCAAGAGCCATTTTCCC 3120

Qy 3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAGGAGAGCGCTTGGGCGTGTGGGGACAC 3180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAGGAGAGCGCTTGGGCGTGTGGGGACAC 3180

Qy 3181 GGTGATGGTTGCCCGGTGTTGCGGCCGTCTCGGCGGCCGCACCTTTGTTGTTTCGGCAGGGTTGGCTAT 3240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3181 GGTGATGGTTGCCCGGTGTTGCGGCCGTCTCGGCGGCCGCACCTTTGTTGTTTGTGACAGGGTTGGCTAT 3240

Qy 3241 GCCCCAGATGGGTGGGCCATTACCGCACCTTTTACGCTGCAGTGTCTCTCTGAACGTGG 3300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3241 GCGCCAGATGGGTGGGCCATTACCGCACCTTTTACGCTGCAGTGTCTCTCTGAACGTGG 3300

Qy 3301 CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC 3360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3301 CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC 3360

Qy 3361 TATCTTCAGATTAGGATCCTGCGCCACTAGCTACATGGATTTGTTGTTGTGACAACGTGTT 3420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3361 TATCTTCAGATTAGGATCCTGCGCCACTAGCTACATGGATTTGTTGTTGTGACAACGTGTT 3420

Qy 3421 GTATACTGCTCACCATGCGAGCAGCAGGGCGCCGGTTGGGCTCATCCACAGGCTCTATACA 3480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3421 GTATACTGCTCACCATGCGAGCAGCAGGGCGCCGGTTGGGCTCATCCACAGGCTCTATACA 3480

Qy 3481 CCCAATAACCGTTGACGGCGGCTAATGACCAGGACATCATCAACACCACCATGTGAGCTGG 3540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3481 CCCAATAACCGTTGACGGCGGCTAATGACCAGGACATCATCAACACCACCATGTGAGCTGG 3540

Qy 3541 GTCCCTTACTGGTGCTCTTGCGGGGAGACCAAGGGTATCTGGTAACACGACTGGGGTC 3600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3541 GTCCCTTACTGGTGCTCTTGCGGGGAGACCAAGGGTATCTGGTAACACGACTGGGGTC 3600

Qy 3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGGGGGCCCTTCCCAT 3660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGGGGGCCCTTCCCAT 3660

Qy 3661 GGCTGTTGCCAAGGGTTCTTCAGGTGCCCCGATTCTGTTGCTCGTTCGTGTGCATGTTATTGG 3720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3661 GGCTGTTGCCAAGGGTTCTTCAGGTGCCCCGATTCTGTTGCTCGTTCGTGTGCATGTTATTGG 3720

Qy 3721 GATGTTCACCCGCTGCTAGAAATTCTGGGGTCAGTCAGTCACAGCACATGCCACTCTTGATACAAAACCTAC 3780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3721 GATGTTCACCCGCTGCTAGAAATTCTGGGGTCAGTCAGTCACAGCACATGCCACTCTTGATACAAAACCTAC 3780

Qy 3781 GGTGTGTGCTGGATCATCCCCAGTACACAGCACATGCCCCCACTGGCAGTGCAAGTCAAC 3840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3781 GGTGTGTGCTGGATCATCCCCAGTACACAGCACATGCCCCCACTGGCAGTGCAAGTCAAC 3840

Qy 3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCTTGTTCGTTATTGG 3900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCTTGTTCGTTATTGG 3900

Qy 3901 CAAATTACCACTTTCCTTACACGAGGAAGTATGAGGGTCTTGGTCCTAAATCCCAGTGT 3960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3901 CAAATTACCACTTTCCTTACACGAGGAAGTATGAGGGTCTTGGTCCTAAATCCCAGTGT 3960

Qy 3961 GGCTACAACAGCATCAATGCCAAAGTACACGCGACGTACGGCGGTGAATCCAAATTG 4020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3961 GGCTACAACAGCATCAATGCCAAAGTACACGCGACGTACGGCGGTGAATCCAAATTG 4020

Qy 4021 CTATTTTAATGGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGGCAT 4080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4021 CTATTTTAATGGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGGCAT 4080

Qy 4081 GTACCTGACCGGAGCAGTGTCCCGGAACTATGATGTAATCATTGTGACGAATGCCATGC 4140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4081 GTACCTGACCGGAGCAGTGTCCCGGAACTATGATGTAATCATTGTGACGAATGCCATGC 4140

Qy 4141 TACCGATGCAACCACCGTGTTGGCATTGGCAGAAAGGTCCTCAACCGAAGCTCCATCCAAAAA 4200
Db 4141 TACCGATGCAACCACCGTGTTGGCATTGGCAGAAAGGTCCTCAACCGAAGCTCCATCCAAAAA 4200
```

Left column (positions 4141–5280):

```
Db  4141  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TACCGGATGCAACACCACCGTGGTTGGGCATTGCCACGGCTACCCCCCTGGAGTAATCCTACCACCACATGC  4200

Qy  4201  TGTTAGGCTAGTGGTTCTTGCCACGGCTACCGGCTACCCCCCTGGAGTAATCCTACCACCACATGC  4260
Db  4201  TGTTAGGCTAGTGGTTCTTGCCACGGCTACCGGCTACCCCCCTGGAGTAATCCTACCACCACATGC  4260

Qy  4261  CAACATAACTGAGATTCAATTAACCGATGAGGCACTATCCCGTTTCATGGAAAAAAGAT  4320
Db  4261  CAACATAACTGAGATTCAATTAACCGATGAGGCACTATCCCGTTTCATGGAAAAAAGAT  4320

Qy  4321  TAAGGAGGAAAAATCGAAGAAAGGGAGACACTTATCTTTGAGGGCTACCAAAAAACACTG  4380
Db  4321  TAAGGAGGAAAAATCGAAGAAAGGGAGACACTTATCTTTGAGGGCTACCAAAAAACACTG  4380

Qy  4381  TGATGAGCTTGCTAACGAGTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG  4440
Db  4381  TGATGAGCTTGCTAACGAGTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG  4440

Qy  4441  ATGTGACATCTCAAAAATCCCTGAGGCGACTGTGTAGTAGTTGCGAGCCTCATGATGCCTTGTG  4500
Db  4441  ATGTGACATCTCAAAAATCCCTGAGGCGACTGTGTAGTAGTTGCGAGCCTCATGATGCCTTGTG  4500

Qy  4501  TACAGGGTACAACTGGTGACTTTGACCCTACTTTCACCATGGTGTTCGTGTGCGGGGTTTCAGC  4560
Db  4501  TACAGGGTACAACTGGTGACTTTGACCCTACTTTCACCATGGTGTTCGTGTGCGGGGTTTCAGC  4560

Qy  4561  ATGCCATGGTGACCTTGACCCTACTTTGTCTTGCTTGACCCTACTTTCGTGTGCGGGGTTTCAGC  4620
Db  4561  ATGCCATGGTGACCTTGACCCTACTTTGTCTTGCTTGACCCTACTTTCGTGTGCGGGGTTTCAGC  4620

Qy  4621  AATAGTTAAAGGCCAGCGTAGGGCCCACAGGCCGTGGGAGAGAGTCGGCATATACTACTA  4680
Db  4621  AATAGTTAAAGGCCAGCGTAGGGCCCACAGGCCGTGGGAGAGAGTCGGCATATACTACTA  4680

Qy  4681  TGTAGACGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGTTGTTCCTGAATGTTGAAGCCTT  4740
Db  4681  TGTAGACGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGTTGTTCCTGAATGTTGAAGCCTT  4740

Qy  4741  TTGTGTCAATGGCATGGCCAAGGCATGGTATGGTTTGTCATCAACAGAGCTCAAACTATTCTGACAC  4800
Db  4741  TTGTGTCAATGGCATGGCCAAGGCATGGTATGGTTTGTCATCAACAGAGCTCAAACTATTCTGACAC  4800

Qy  4801  CTATCGCACCCAACCTGGGTTACCTGGGTTACCGAGTGGAGTGGCTGATCT  4860
Db  4801  CTATCGCACCCAACCTGGGTTACCTGGGTTACCGAGTGGAGTGGCTGATCT  4860

Qy  4861  CTTTTCTCATGGTCAACCCGACCTTCATTTGTCAATATGTTGAAGAGACTGCTGACAA  4920
Db  4861  CTTTTCTCATGGTCAACCCGACCTTCATTTGTCAATATGTTGAAGAGACTGCTGACAA  4920

Qy  4921  CGACGCAGCCAAGGCATGGCCAAGGCCAAGGCCCAGCGCCCAGGCCCACACTGGTATGCTGCTCC  4980
Db  4921  CGACGCAGCCAAGGCATGGCCAAGGCCAAGGCCCAGCGCCCAGGCCCACACTGGTATGCTGCTCC  4980

Qy  4981  CAATGACGCACCACGCGGCGCTTGGCAGGCAGCCCGCTTGTGTGGGGTTCTGTG  5040
Db  4981  CAATGACGCACCACGCGGCGCTTGGCAGGCAGCCCGCTTGTGTGGGGTTCTGTG  5040

Qy  5041  GCGCTTGGACGGCGCTGACGCGCTGGCCTGGCCCAGAGCCCGCCACTCGCGTTGCGGTTGACCAGATACCA  5100
Db  5041  GCGCTTGGACGGCGCTGACGCGCTGGCCTGGCCCAGAGCCCGCCACTCGCGTTGCGGTTGACCAGATACCA  5100

Qy  5101  AATGTGCTTCACTGAAGTCAATAACTTCTCGGGGACAGTTGTCATGGCTATGCTGCGGAGT  5160
Db  5101  AATGTGCTTCACTGAAGTCAATAACTTCTCGGGGACAGTTGTCATGGCTATGCTGCGGAGT  5160

Qy  5161  GGCTATGGCTTATCTAGCACTTGACACTTTTGCGCGCCACTTGTGTGCGGCGGTTGCTGGTC  5220
Db  5161  GGCTATGGCTTATCTAGCACTTGACACTTTTGCGCGCCACTTGTGTGCGGCGGTTGCTGGTC  5220

Qy  5221  TATTACACATCAGTCCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAGAAATCGT  5280
Db  5221  TATTACACATCAGTCCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAGAAATCGT  5280
```

Right column (positions 5221–6360):

```
Db  5221  TATTGCATCAGTCCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAGAAATCGT  5280

Qy  5281  GGAGGAGTGTGCATCATTCCCTGGAGGCCATGTGTTGCTGCAATTGACAAGCTGAA  5340
Db  5281  GGAGGAGTGTGCATCATTCCCTGGAGGCCATGTGTTGCTGCAATCGATAAGCTGAA  5340

Qy  5341  GAGTACAATCACCACAACTAGTCCTTCACATTGGAAAACTTAACAC  5400
Db  5341  GAGTACAATCACCACAACTAGTCCTTCACATTGGAAAACTTAACAC  5400

Qy  5401  CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGTTTAGT  5460
Db  5401  CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGTTTAGT  5460

Qy  5461  CACTTTACCTGACAATCCTTGTCTTTGCTTGCTTTCATTGGGGCGCAATTGCGTCCAA  5520
Db  5461  CACTTTACCTGACAATCCTTGTCTTTGCTTGCTTTCATTGGGGCGCAATTGCGTCCAA  5520

Qy  5521  CCCACTACCTCACAAGATCAAAATGTTCCTGTCACTTATTTGGAGGCGCACTGGGGAACAGCTCT  5580
Db  5521  CCCACTACCTCACAAGATCAAAATGTTCCTGTCACTTATTTGGAGGCGCACTGGGGAACAGCTCT  5580

Qy  5581  GCTTACAGACGCTAGAGGCGCACTGGCGTTGTCTTGACATGGCGTATGCTGCGCCCTC  5640
Db  5581  GCTTACAGACGCTAGAGGCGCACTGGCGTTGTCTTGACATGGCGTATGCTGCGCCCTC  5640

Qy  5641  TGGTACATGGACATCGGTGGGGTTTGTCTTTGACATCGGTGGGGTTTCAGC  5700
Db  5641  TGGTACATGGACATCGGTGGGGTTTGTCTTTGACATCGGTGGGGTTTCAGC  5700

Qy  5701  ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGGGCCCACTATGATCAGCT  5760
Db  5701  ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGGGCCCACTATGATCAGCT  5760

Qy  5761  TGCTGGTTAGTCTACTCCGGTCAATCCGGTCGTGTGGGCGTCTTGTCAGC  5820
Db  5761  TGCTGGTTAGTCTACTCCGGTCAATCCGGTCGTGTGGGCGTCTTGTCAGC  5820

Qy  5821  TTGTGCAATGTTGCTTTGACAACAGCAGGGCCAGATCACTGGCCCAACAGACTTCTTAC  5880
Db  5821  TTGTGCAATGTTGCTTTGACAACAGCAGGGCCAGATCACTGGCCCAACAGACTTCTTAC  5880

Qy  5881  TATGCTTGCTGTAGGAGCAACACTGTATGTAATAGTACTTTATTGCCACTCGTGACATCCG  5940
Db  5881  TATGCTTGCTGTAGGAGCAACACTGTATGTAATAGTACTTTATTGCCACTCGTGACATCCG  5940

Qy  5941  CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCCTGGAGTGTCATATCAGCTTGCAT  6000
Db  5941  CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCCTGGAGTGTCATATCAGCTTGCAT  6000

Qy  6001  CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGCGGCCTCATTGCTCTTAGAGAT  6060
Db  6001  CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGCGGCCTCATTGCTCTTAGAGAT  6060

Qy  6061  TTGGCAGTATGTGTGCAATTTGTCTTTGTGATTTGCTTTAAAGTCCTTAAAGCTGGAGTTCA  6120
Db  6061  TTGGCAGTATGTGTGCAATTTGTCTTTGTGATTTGCTTTAAAGTCCTTAAAGCTGGAGTTCA  6120

Qy  6121  GAGCATGGTTAACATTCCTGGTTGTCCTTTCTTTGTGTTCTTCACAGCGTGCCAGAAGGGGTACAAGGGCCC  6180
Db  6121  GAGCATGGTTAACATTCCTGGTTGTCCTTTCTTTGTGTTCTTCACAGCGTGCCAGAAGGGGTACAAGGGCCC  6180

Qy  6181  CTGGATTGGATCAGGTATGCTCCAAGCACTCGCGTGCTCGAACTCATCTTTTC  6240
Db  6181  CTGGATTGGATCAGGTATGCTCCAAGCACTCGCGTGCTCGAACTCATCTTTTC  6240

Qy  6241  TGTTGAGAATGGTTTGCAAAACTTTACAAAAGGACCCAGAAACTTGTTCAAATTACTGGAG  6300
Db  6241  TGTTGAGAATGGTTTGCAAAACTTTACAAAAGGACCCAGAAACTTGTTCAAATTACTGGAG  6300

Qy  6301  AGGGGCTGTTCCAGTCAACGCGTGTCCAGTCAACGCGTGTCGGCTGTCGGTCGGCTGAGACCGACCCAACTGATTG  6360
Db  6301  AGGGGCTGTTCCAGTCAACGCGTGTCCAGTCAACGCGTGTCGGCTGTCGGTCGGCTGAGACCGACCCAACTGATTG  6360
```

Right column (positions 7441–8580):

```
Qy  7441 GTTCCCCCATCATACCACAAGCAAGCAGATTGAGAGTTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441 GTTCCCCCATCATACCACAAGCAAGCAGATTGAGAGTTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500

Qy  7501 CGGTGTCATGTGGCATATGATGAAGTAGCAGCTCACACGCCCTCTCAAGTCTGCTAAGTC 7560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7501 CGGTGTCATGTGGCATATGATGAAGTAGCAGCTCACACGCCCTCTCAAGTCTGCTAAGTC 7560

Qy  7561 CCACATCACTGGCCTTCGGGGCACTGATCGTTCGTTCTCTGAGCAGCCCGCAAGGCGTGTTCT 7620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7561 CCACATCACTGGCCTTCGGGGCACTGATCGTTCGTTCTCTGAGCAGCCCGCAAGGCGTGTTCT 7620

Qy  7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680

Qy  7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAAACCAACAAAGAAACCCCCAAG 7740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAAACCAACAAAGAAACCCCCAAG 7740

Qy  7741 GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGT 7800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7741 GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGT 7800

Qy  7801 TGCTCCTGACGTAGTTGCATGGGAGATGCGTACGGGTTTGTAGATCCACGTAC 7860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7801 TGCTCCTGACGTAGTTGCATGGGAGATGCGTACGGGTTTGTAGATCCACGTAC 7860

Qy  7861 CCGTGTCAAGCGTCTGTTGTGCGATGTGGTCACCCGATGCGAGCCACATGCGGATAC 7920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7861 CCGTGTCAAGCGTCTGTTGTGCGATGTGGTCACCCGATGCGAGCCACATGCGGATAC 7920

Qy  7921 AGTGTGTTTGACAGTACCATCACACCGAGGATATCGTGGCTGCTGACCTGCTGAAGGTAAATGC 7980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7921 AGTGTGTTTGACAGTACCATCACACCGAGGATATCGTGGCTGCTGACCTGCTGAAGGTAAATGC 7980

Qy  7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCATTCACACCATTGCGAGGCAGTTATA 8040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCATTCACACCATTGCGAGGCAGTTATA 8040

Qy  8041 CGTTGGAGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTGTAGGTGTAGGTC 8100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8041 CGTTGGAGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTGTAGGTGTAGGTC 8100

Qy  8101 TTCCGGCGTCTATACTACCTCAAGTTGACCTGCTGGCTGAAGGTAAATGC 8160
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8101 TTCCGGCGTCTATACTACCTCAAGTTGACCTGCTGGCTGAAGGTAAATGC 8160

Qy  8161 TGCAGCCGAACAGGCTGGCATGGAAGAACCCTCGCTTCCTTATTTGCGGCCGATGATTGCAC 8220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161 TGCAGCCGAACAGGCTGGCATGGAAGAACCCTCGCTTCCTTATTTGCGGCCGATGATTGCAC 8220

Qy  8221 CGTAATTTGGAAGAGCGCCGAGCAATGCGTGTCTTTGCTAG 8280
         |||||||||||||||||||||||||||||||||||||||||
Db  8221 CGTAATTTGGAAGAGCGCCGAGCAATGCGTGTCTTTGCTAG 8280

Qy  8281 CTTGGATGAAGGTGATGGGTGCACCACAAGATTGTGCGCTCAACCCCAAATACAGTTTGGA 8340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281 CTTGGATGAAGGTGATGGGTGCACCACAAGATTGTGCGCTCAACCCCAAATACAGTTTGGA 8340

Qy  8341 AGAATTAACATCATGTCTCATCAAATGTTACCTCGGAATTACCAAAAGTGGCAAGCCTTA 8400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8341 AGAATTAACATCATGTCTCATCAAATGTTACCTCGGAATTACCAAAAGTGGCAAGCCCTTA 8400

Qy  8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCGCCGAGGGGTCTGGG 8460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCGCCGAGGGGTCTGGG 8460

Qy  8461 ATACAACCCAGTGCGTGTTGGCGTGTCCATTTCATGGAGCAGAGTGCTCTTTCATGGAGCAAACTTCCCGA 8520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8461 ATACAACCCAGTGCGTGTTGGCGTGTCCATTTCATGGAGCAGAGTGCTCTTTCATGGAGCAAACTTCCCGA 8580
```

Left column (positions 6361–7440):

```
Qy  6361 GACTAGTCTTGTCGTCAATTATGGCGTTAGGACTACTGTAAATATGAGAAATGGGAGA 6420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361 GACTAGTCTTGTCGTCAATTATGGCGTTAGGACTACTGTAAATATGAGAAATGGGAGA 6420

Qy  6421 TCACATTTTGTTACAGCAGTATCGTCTCCCAAATGTCGTTTCACCCAGGTGCCCCCAAC 6480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421 TCACATTTTGTTACAGCAGTATCGTCTCCCAAATGTCGTTTCACCCAGGTGCCCCCAAC 6480

Qy  6481 CTTGAGAGCTGCAGTGGCCGTGACGGGCTACAGGTTATCTGTAGGTGTTCAGTTCAGGTTATCTAGGTGAGCCCA 6540
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481 CTTGAGAGCTGCAGTGGCCGTGACGGGCTACAGGTTATCTGTAGGTGTTCAGTTCAGGTTATCTAGGTGAGCCCA 6540

Qy  6541 AACTCCTTGGACGACATCTGCTTGCTGTGTTACGGTCCTGACGGTAAGGGTAAAACTGTTAA 6600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541 AACTCCTTGGACGACATCTGCTTGCTGTGTTACGGTCCTGACGGTAAGGGTAAAACTGTTAA 6600

Qy  6601 GCTTCCCTCCCGCGGTTGACGGTCACCACCACCTGGTTGCGCATGCGCATGCAACTTAATTTGCGTGA 6660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601 GCTTCCCTCCCGCGGTTGACGGTCACCACCACCTGGTTGCGCATGCGCATGCAACTTAATTTGCGTGA 6660

Qy  6661 TGCACTTGAGACAAATGACTGTAATTCACAAACAACACTCCTAGTGATGAAGCCGCAGT 6720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661 TGCACTTGAGACAAATGACTGTAATTCACAAACAACACTCCTAGTGATGAAGCCGCAGT 6720

Qy  6721 GTCCGCTCTTGTTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC 6780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721 GTCCGCTCTTGTTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC 6780

Qy  6781 AGGAGTGTCATGTCCTGAAAAGCTCGAAGCTCCTGAAGAGAGTAGTGGTAAGAAA 6840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781 AGGAGTGTCATGTCCTGAAAAGCTCGAAGCTCCTGAAGAGAGTAGTGGTAAGAAA 6840

Qy  6841 GCGCCAGTTCCGGGCAAGAACTGTTCGTTCGTTCCTGCCTTGCCTTGCTTACCTGTTGCCTTACCTTCCGAGATCCGTCCC 6900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841 GCGCCAGTTCCGGGCAAGAACTGTTCGTTCGTTCCTGCCTTGCCTTGCTTACCTGTTGCCTTACCTTCCGAGATCCGTCCC 6900

Qy  6901 AGGAGTGTCATGTCCTGAAAAGCTCCGACGAAGTAGAGTCCTTCAAACCT 6960
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901 AGGAGTGTCATGTCCTGAAAAGCTCCGACGAAGTAGAGTCCTTCAAACCT 6960

Qy  6961 CCCTCCTTCACCACCTGTTCTCACCAGTTGGCCATGCCATGGATGCCGAGGCGGGTGA 7020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961 CCCTCCTTCACCACCTGTTCTCACCAGTTGGCCATGCCATGGATGCCGAGGCGGGTGA 7020

Qy  7021 GTGTAACCCTTTCACTGCGCAATTGCAATTGGATGTGCAAACAGGCGAAATGGTCTCTGAATGGTCAGACGTAGACGTCTCTGAATGGTCTGATGA 7080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021 GTGTAACCCTTTCACTGCGCAATTGCAATTGGATGTGCAAACAGGCGAAATGGTCTCTGAATGGTCAGACGTAGACGTCTCTGAATGGTCTGATGA 7080

Qy  7081 TTTACCCAGTTACCCTCCCAAAAAGGACGTCTCTGAATGGTCGTTCTTCAAAACTGCTTCAAAACTGCTTCTAAAGT 7140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081 TTTACCCAGTTACCCTCCCAAAAAGGACGTCTCTGAATGGTCGTTCTTCAAAACTGCTTCAAAACTGCTTCTAAAGT 7140

Qy  7141 GGCTACAACCGCTTCCAGCTACTGGCCTCGGCCACCACTAGTGGTTCCTCAAACGAAGATACGGGGAAAGGA 7200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141 GGCTACAACCGCTTCCAGCTACTGGCCTCGGCCACCACTAGTGGTTCCTCAAACGAAGATACGGGGAAAGGA 7200

Qy  7201 TTCCACTCAGTCAGCAGCCCGCCCCCCGCCAAACGGCCTACAAAAAGAGAGTGGGAAAGAGTGAGTT 7260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201 TTCCACTCAGTCAGCAGCCCGCCCCCCGCCAAACGGCCTACAAAAAGAGAGTGGGAAAGAGTGAGTT 7260

Qy  7261 TTTCGTGCAGCATGAGCTACACCGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261 TTTCGTGCAGCATGAGCTACACCGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320

Qy  7321 TCTGTGTCTGCAACTGCGGGGATGCGGGGCCATCAACTAGTGGTTCTTAATAATGGTGTATGT 7380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7321 TCTGTGTCTGCAACTGCGGGGATGCGGGGCCATCAACTAGTGGTTCTTAATAATGGTGTATGT 7380

Qy  7381 GACTGAGCGCGGGGATGCGGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCT 7440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381 GACTGAGCGCGGGGATGCGGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCT 7440
```

OS   Hepatitis GB virus B.
XX
PN   US2004039187-A1.
XX
PD   26-FEB-2004.
XX
PF   03-JUL-2002; 2002US-00189359.
XX
PR   04-JUN-1999;   99US-0137665P.
PR   05-JUN-2000;   2000US-00587653.
XX
PA   (TEXA) UNIV TEXAS SYSTEM.
PA   (INSP) INST PASTEUR.
XX
PI   Martin A, Sangar DV, Lemon SM, Rijnbrand R;
XX
DR   WPI; 2004-203294/19.
XX
PT   New GB virus B and/or hepatitis C virus (HCV) sequences, useful in
PT   diagnosing and in treating HCV and in investigating the mechanisms for
PT   the different biological properties of the viruses.
XX
PS   Claim 10; SEQ ID NO 2; 58pp; English.
XX
CC   The invention describes a new isolated polynucleotide (I) encoding a 3'
CC   sequence of the GB virus B (GBV-B) genome, or which comprises a chimeric
CC   GBV-B genome, where at least part, but not all of a 5' nontranslated
CC   region (NTR) sequence is derived from a hepatitis C virus (HCV) 5' NTR.
CC   (I) is a GB virus B and/or hepatitis C virus polynucleotide comprising a
CC   fully defined of 260 or 9399 bp (SEQ ID NOS: 1 or 2). The polynucleotides
CC   or chimaeras are useful diagnosing or treating hepatitis C virus (HCV)
CC   and in investigating the mechanisms for the different biological
CC   properties of the viruses. This sequence represents a Hepatitis GB virus
CC   B (GBV-B) 3' terminal polynucleotide.
XX
SQ   Sequence 9399 BP; 2124 A; 2322 C; 2432 G; 2521 T; 0 U; 0 Other;

Query Match   99.9%; Score 9386.2; DB 12; Length 9399;
Best Local Similarity   99.9%; Pred. No. 0;
Matches 9391; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy   1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGAATGCTCCTGGAGCCCACCTATATGT 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGAATGCTCCTGGAGCCCACCTATATGT 60

Qy   61 CAGGGCGTGGGGGATTCCCCGTCTCGCAGAAGGGTGAGCCAACCACCTTAGTAT 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CAGGGCGTGGGGGATTCCCCGTCTCGCAGAAGGGTGAGCCAACCACCTTAGTAT 120

Qy   121 GTAGGCGCGGGGACTCATGACCGTCGGGTGATGACAAGCGCAAGCTTGACTTGGATGGC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTAGGCGCGGGGACTCATGACCGTCGGGTGATGACAAGCGCAAGCTTGACTTGGATGGC 180

Qy   181 CCTGATGGGCGTTCATGGGTTCGGTGGGTGGTGCGGCGTTAGGCAGCCTCCACGCCACCA 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCTGATGGGCGTTCATGGGTTCGGTGGGTGGTGCGGCGTTAGGCAGCCTCCACGCCACCA 240

Qy   241 CCTCCCAGATAGAGCGGCGGCCACTGTAGGGAAGACCGGGAGAACCGGGGCACTACCAAGGACG 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCCCAGATAGAGCGGCGGCCACTGTAGGGAAGACCGGGAGAACCGGGGCACTACCAAGGACG 300

Qy   301 CAGACCTCTTTTGAGTATCACGCCTCCGGAAGCCACCTATATGTGT 360
            ||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGACCTCTTTTGAGTATCACGCCTCCGGAAGCCACCTATATGTGT 360

Qy   361 TGGGATGGTTGGGGTTAGCCATCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TGGGATGGTTGGGGTTAGCCATCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT 420

Qy   421 CTGGAGTCTCTGTAGACCGTAGCACATGCCTGTATTTCTACTCAAACAAGTCCTGTACC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTGGAGTCTCTGTAGACCGTAGCACATGCCTGTATTTCTACTCAAACAAGTCCTGTACC 480

Db   8521 TAGCCGTGTGTTGGCGTGTCCAATTCATGCAGGATGCGTCTCTTGAGGAGCAAAACTTCCCGA 8580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   8581 GACTGTGACCTTGACTGGTATGGGAGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8581 GACGGTGACCTTGACTGGTATGGGAGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640

Qy   8641 CATCATTGCTGGTGTGCACGGTATTGAGGCTTCTCCGGTGCGCTACCAACGCTGA 8700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8641 CATCATTGCTGGTGTGCACGGTATTGAGGCTTCTCCGGTGCGCTACCAACGCTGA 8700

Qy   8701 GATCCTCAGAGTTCCCAATCACTAACAGACATGACAATGACCACGTGCCCCCGCGAGCCTGGCG 8760
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8701 GATCCTCAGAGTTCCCAATCACTAACAGACATGACAATGACCACGTGCCCCCGCGAGCCTGGCG 8760

Qy   8761 AAAGAAAGCCAGGGGCGTCCGCCAGCCGCCAAGAGGCGTGGCGGAGCACACGCAAAATT 8820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8761 AAAGAAAGCCAGGGGCGTCCGCCAGCCGCCAAGAGGCGTGGCGGAGCACACGCAAAATT 8820

Qy   8821 GGCTCGCTTCCTTCCTCTGGCATGCTACATCTAGACCTTCTACCAGATTGGATAAGACGAG 8880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8821 GGCTCGCTTCCTTCCTCTGGCATGCTACATCTAGACCTTCTACCAGATTGGATAAGACGAG 8880

Qy   8881 CGTGGCTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGAGGGGGATGTGTT 8940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8881 CGTGGCTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGAGGGGGATGTGTT 8940

Qy   8941 TATTACACCAGAAGAGATTGCAGAAGTTCCTTGTGAAGTATTGGCGTGTCATTGTTT 9000
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8941 TATTACACCAGAAGAGATTGCAGAAGTTTGCAGAAGTTCCTTGTGAAGTATTGGCGTGTCATTGTTT 9000

Qy   9001 TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9001 TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060

Qy   9061 CTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCCGGGC 9120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9061 CTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCCGGGC 9120

Qy   9121 TTAACGACCCCGCCGATGTGAGTTGGCGACCATGGTTCGGTGAA 9180
            |||||||||||||||||||||||||||||||||||||||||||
Db   9121 TTAACGACCCCGCCGATGTGAGTTGGCGACCATGGTTCGGTGAA 9180

Qy   9181 GCCATGGTCTGAAGGGGATGACGTCCCCTTCGGCTCATCCACAAAAACCGTCTCGGGTGG 9240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9181 GCCATGGTCTGAAGGGGATGACGTCCCCTTCGGCTCATCCACAAAAACCGTCTCGGGTGG 9240

Qy   9241 GTGAGGAGTCCTGGCGTGTGTGGGAAGCAGTCAGTATAATTCCCGTCGTGTGTGGTGACGC 9300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9241 GTGAGGAGTCCTGGCGTGTGTGGGAAGCAGTCAGTATAATTCCCGTCGTGTGTGGTGACGC 9300

Qy   9301 CTCACGACGTATTTGTCCGCGTGTGCAGAGCGGCTGCAGACCCGGTTTTTG 9360
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   9301 CTCACGACGTATTTGTCCGCGTGTGCAGAGCGGCTGCAGACCCGGTTTTTG 9360

Qy   9361 TTCCAAGCGGAGGGCAACCCCCGCCGCTTGGAATTAAAAACT 9399
            ||||||||||||||||||||||||||||||||||||||||||
Db   9361 TTCCAAGCGGAGGGCAACCCCCGCCGCTTGGAATTAAAAACT 9399

RESULT 3
ADJ64244
ID   ADJ64244 standard; DNA; 9399 BP.
XX
AC   ADJ64244;
XX
DT   20-MAY-2004 (first entry)
XX
DE   GB virus B 3' terminal polynucleotide fragment seqid 2.
XX
KW   antiinflammatory; hepatotropic; virucide; GB virus B; GBV-B;
KW   hepatitis C virus; HCV; 3' terminal; ds.
XX

Left column:

```
Qy   481 TGCGCCCAGAACGCCAAGAACAAGCAGACGCAGCGCAGGCTTCATATCCTGTGTCCATTAAAAC 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TGCGCCCAGAACGCCAAGAACAAGCAGACGCAGCGCAGGCTTCATATCCTGTGTCCATTAAAAC 540

Qy   541 ATCTGTTGAAAGGGGACAACGAGCAAAGCGCAAAGTCCAGCGCGATGCTCGGCCTCGTAA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 ATCTGTTGAAAGGGGACAACGAGCAAAGCGCAAAGTCCAGCGCGATGCTCGGCCTCGTAA 600

Qy   601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC 660

Qy   661 TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTCGCAATCTTGGAATCCTTCTCGA 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTCGCAATCTTGGAATCCTTCTCTGA 720

Qy   721 TTACCCTTTGGGGTGATTGGTACAACTCACACACCTCACACACCTTCACACACCTTCGGCCTG 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TTACCCTTTGGGGTGATTGGTACAACTCACACACCTCACACACCTTCGGCCTG 780

Qy   781 GGCAGGAGGCGGTCGTTGCACCAGTCCTGCTTGCTGCAGATACGCTTGCTGGGAGTGCAA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GGCAGGAGGCGGTCGTTGCACCAGTCCTGCTTGCTGCAGATACGCTTGCTGGGAGTGCAA 840

Qy   841 CTGGGCTACTGTTGGTTCGGTGTCCACCTTTTGTGTATGTCTGCTATCTTTGGCCTG 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CTGGGCTACTGTTGGTTCGGTGTCCACCTTTTGTGTATGTCTGCTATCTTTGGCCTG 900

Qy   901 TCCCTGTAGTGGGGCGCGGGTCACTGACCCAGTCCTTGCTGCAGACACCACAATTG 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCCTGTAGTGGGGCGCGGGTCACTGACCCAGTCCTTGCTGCAGACACCACAATTG 960

Qy   961 CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCTTCCACTTGCCTACACGAGCCTGGTTG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCTTCCACTTGCCTACACGAGCCTGGTTG 1020

Qy  1021 TGTGATCTGTCGGACGAGTGCTGGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 TGTGATCTGTCGGACGAGTGCTGGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA 1080

Qy  1081 TTGGACTGGCACGGACTCCTTCTTTGGCTGACCACATTGATTTTGTTATGGGGCTCTTGT 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TTGGACTGGCACGGACTCCTTCTTTGGCTGACCACATTGATTTTGTTATGGGGCTCTTGT 1140

Qy  1141 GACCTGTGACGCCCTTCGAGTTTCGCGTGTCGTGTGTATTAGTCGGTGACTG 1200
         |||||||||||||||||||||||||||||||||||||||||||||
Db  1141 GACCTGTGACGCCCTTCGAGTTTCGCGTGTCGTGTGTATTAGTCGGTGACTG 1200

Qy  1201 GCTTGTCAGGCACTGGCTTATTCACATGAAACTCAATGAAACTGGTACTTGTTACCTGGA 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GCTTGTCAGGCACTGGCTTATTCACATGAAACTCAATGAAACTGGTACTTGTTACCTGGA 1260

Qy  1261 AGTGGCCACTGGTACACTACCTGGGTTATCGGGTTCCTAGGGTTATCGGGTGGATGGCCGGCAAGGT 1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 AGTGGCCACTGGTACACTACCTGGGTTATCGGGTTCCTAGGGTTATCGGGTGGATGGCCGGCAAGGT 1320

Qy  1321 CGAGGGTATCAGTGCTCATATAGCTGAGTTTCGCCCTTTGATGATACCATGTCCGACTAT 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CGAGGGTATCAGTGCTCATATAGCTGAGTTTCGCCCTTTGATGATACCATGTCCGACTAT 1380

Qy  1381 GTTTAGCAGTGTACACTACCTGCGGCTTGGCGCTCTGCGGCTTATGCCTCTCGGGGCAA 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 GTTTAGCAGTGTACACTACCTGCGGCTTGGCGCTCTGCGGCTTATGCCTCTCGGGGCAA 1440

Qy  1441 GTGTGTTCAGTCAGTGCTTATACAGAAGGCCACCTTCTGAGTTTCGCGCCCTTGAAAACCCCAT 1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 GTGGTATCAGTCAGTGCTTATACAGAAGCGACCTTCTGAGTTTCGCGCCCTTGAAAACCCTAT 1500

Qy  1501 CAGGGTGCCCACTGCTCAATAGCTGAGTTTTGCTCGCCCTTTGATGATACCATGTCC 1560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 CAGGGTGCCCACTGCTCGATGCTCAATAGCTGAGTTTTGCTCGCCCTTTGATGATACCATGTCC 1560

Qy  1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTTGTTACAGTGGAC 1620
```

Right column:

```
Db  1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTTGTTACAGTGTACAGTGGAC 1620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1621 CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTCGTTGTGTACCCCTATACAATCCCTG 1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTCGTTGTGTACCCCTATACAATCCCTG 1680

Qy  1681 TGCGAGGGGATGTATGGTTAAATTCAAAATAACACATGGGGTTGCTGCGTATTCGCAA 1740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 TGCGAGGGGATGTATGGTTAAATTCAAAATAACACATGGGGTTGCTGCGTATTCGCAA 1740

Qy  1741 TGTGCCATCGTACTGCCACTATGGCCACTGATGGCAGTGTGAACGACACTCGCAACACTTA 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 TGTGCCATCGTACTGCCACTATGGCCACTGATGGCAGTGTGAACGACACTCGCAACACTTA 1800

Qy  1801 CGAAGCATGCGGTGTAACACCATGCCTAACACCGCATGGCACAACGGCTCAGCCCTGAA 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 CGAAGTATGCGGTGTAACACCATGCCTAACACCGCATGGCACAACGGCTCAGCCCTGAA 1860

Qy  1861 ATTGGCTATATACAATACCCGGTCTAAAGAAAATGTTAAACCTCATAATTGGATGTC 1920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 ATTGGCTATATTACAATACACCCGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC 1920

Qy  1921 AGGGCCATTTGTATTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA 1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 AGGGCCATTTGTATTTGAGGGATCAGATACCCTATAGTTTACTTTTTACTTTATGACCCTGTGAA 1980

Qy  1981 TTCCACTCTTCCTACCACCCGGAGAGTTGCCGGGTACCCGGTTGCCCCACCTGTGGTACG 2040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 TTCCACTCTTCCTACCACCCGGAGAGTTGCCGGGTACCCGGTTGCCCCACCTGTGGTACG 2040

Qy  2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 TGGTTCTTGGTTACAGGTTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100

Qy  2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC 2160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC 2160

Qy  2161 TTTGTCTCTACGGGAGTTACCACCAAGGCCGTGGTGCTAATTCTTGTGGGGTTGTGTGG 2220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 TTTGTCTCTACGGGAGTTACCACCAAGGCCGTGGTGCTAATTCTTGTGGGGTTGTGTGG 2220

Qy  2221 CAGCAAGTATCTTATTTTAGCCTATTTTGCCTTGTCCCTTTTGTTTGGGCGCGGCCTTC 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 CAGCAAGTATCTTATTTTAGCCTATTTTGCCTTGTCCCTTTTGTTTGGGCGCGGCCTTC 2280

Qy  2281 TGGTTACCCTTGCGTCTGTGCTCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGT 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 TGGTTACCCTTGCGTCTGTGCTCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGT 2340

Qy  2341 TTTGTCTAAAGCTCAAGTAGCTCCTTTTGCTTGCTTTGCTTGTTCATCTGTTGCTATCTCCG 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 TTTGTCTAAAGCTCAAGTAGCTCCTTTTGCTTGCTTTGCTTGTTCATCTGTTGCTATCTCCG 2400

Qy  2401 CTGCAGTATCTGGTTATGCGCTGCCCCATGGCTGCGCGGCTGCTTCCCCCT 2460
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 CTGCAGTATCTGGTTATGCGCTGCCCCATGGCTGCGCGGCTGCTTCCCCCT 2460

Qy  2461 AACTTCTTTGTTGTGCAGCAGCTGCTGCCAACCAGATTATGACTGGTGCGACTGCT 2520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 AACTTCTTTGTTGTGCAGCAGCTGCTGCCAACCAGATTATGACTGGTGCGACTGCT 2520

Qy  2521 AGTGGCAGGGTTAGTTTTGTGGGGCCGGCCGTAACCCGGTCACCGGCCATAGCTCTTGT 2580
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521 AGTGGCAGGGTTAGTTTTGTGGGGCCGGCCGTAACCCGGTCACCGGCCATAGCTCTTGT 2580

Qy  2581 AGTTCCTTGGCCTCTCGGTTAGCGCTTTTAACCCTCTTGCATTTGGTTACCGCGTCTTCAGC 2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581 AGTTCCTTGGCCTCTCGGTTAGCGCTTTTAACCCTCTTGCATTTGGTTACCGCGTCTTCAGC 2640

Qy  2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGTTGTCAT 2700
```

```
Db 3781 GGTGTGTGCTGATACCATCCCAGTACGTACACAGCACATGCCACTCTTGATACAAAACCTAC 3840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 3781 GGTGTGTGCTGATACCATCCCAGTACGTACACAGCACATGCCACTCTTGATACAAAACCTAC 3840

Qy 3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCACTGGCAGCGGCAAGTCAAC 3900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCACTGGCAGCGGCAAGTCAAC 3900

Qy 3901 CAAATTACCACTTTCTTACATGCAGGAGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT 3960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3901 CAAATTACCACTTTCTTACATGCAGGAGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT 3960

Qy 3961 GGCTACAACAGACATCAATGCCAAAGTACATGCACGCGACGTACGGCGTGAATCCAAATTG 4020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3961 GGCTACAACAGACATCAATGCCAAAGTACATGCACGCGACGTACGGCGTGAATCCAAATTG 4020

Qy 4021 CTATTTTAATGCAAATGTACCAACACAGGGGCTTCACTTAGTACGACACATATGGCAT 4080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4021 CTATTTTAATGCAAATGTACCAACACAGGGGCTTCACTTAGTACGACACATATGGCAT 4080

Qy 4081 GTACCTGACCGGACGCATGTCCCGGAACTATGATGTAATCATTGTGACGAATGCCATGC 4140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4081 GTACCTGACCGGACGCATGTCCCGGAACTATGATGTAATCATTGTGACGAATGCCATGC 4140

Qy 4141 TACCGATGCAACCACCGTGTTGGGCATTGGGCCACGGTTACCCCCCGGAGTAATCCCACCACATGC 4200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4141 TACCGATGCAACCACCGTGTTGGGCATTGGGCCACGGTTACCCCCCGGAGTAATCCCACCACATGC 4200

Qy 4201 TGTTAGGCTAGTGGTTCTTGCCACGGTGTGTGGGGACAC 4260
           |||||||||||||||||||||||||||||||||||||||
Db 4201 TGTTAGGCTAGTGGTTCTTGCCACGGTGTGTGGGGACAC 4260

Qy 4261 CAACATAACTGAGATTCAATAACCGATGAAGGCACTATCCCTTTCATGGAAAAAAGAT 4320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4261 CAACATAACTGAGATTCAATAACCGATGAAGGCACTATCCCTTTCATGGAAAAAAGAT 4320

Qy 4321 TAAGGAGGAAAATCTGAAGAAAGGAGGAGACACCTTATCTTTGAGGCTACCAAAAACACTG 4380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4321 TAAGGAGGAAAATCTGAAGAAAGGAGGAGACACCTTATCTTTGAGGCTACCAAAAACACTG 4380

Qy 4381 TGATGAGGTTGCTAACGAGTTAGCTCGAAAGGAATAACAGCTGTCTTCTTACTATAGGGG 4440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4381 TGATGAGGTTGCTAACGAGTTAGCTCGAAAGGAATAACAGCTGTCTTCTTACTATAGGGG 4440

Qy 4441 ATGTGACATCTCAAAAATCCCGAGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4441 ATGTGACATCTCAAAAATCCCGAGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500

Qy 4501 TACAGGGTACACTGGTTGATTCCGTGTATGGGTATGGTTCCGTGTATGGCTAGAAGGCAC 4560
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4501 TACAGGGTACACTGGTTGATTCCGTGTATGGGTATGGTTCCGTGTATGGCTAGAAGGCAC 4560

Qy 4561 ATGCCATGTTGACCCTGACCCTACTTTCACCATGGGGGTGTTCGTGTGTGCGGGGTTTCAGC 4620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4561 ATGCCATGTTGACCCTGACCCTACTTTCACCATGGGGGTGTTCGTGTGTGCGGGGTTTCAGC 4620

Qy 4621 AATAGTTAAAGGCCAGCGGCCGGTCCTGGGAGACTGGCATATACTACTA 4680
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db 4621 AATAGTTAAAGGCCAGCGGCCGGTCCTGGGAGACTGGCATATACTACTA 4680

Qy 4681 TGTAGACGGAGTTGTACCCCTTGGGTATGGGTCCTGAATGCAACATTGTTGAAGCCTT 4740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4681 TGTAGACGGAGTTGTACCCCTTGGGTATGGGTCCTGAATGCAACATTGTTGAAGCCTT 4740

Qy 4741 CGACGCAGCCAAGGCACATGGTATGGTTTGTCATCAACAGAAGCTCAAACTATTCTGGACAC 4800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4741 CGACGCAGCCAAGGCACATGGTATGGTTTGTCATCAACAGAAGCTCAAACTATTCTGGACAC 4800

Qy 4801 CTATCGCACCCAACCTGGTTACCTGCGGATAGGAGCAAATTTGGACGAGTGGGCTGATCT 4860
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4801 CTATCGCACCCAACCTGGTTACCTGCGGATAGGAGCAAATTTGGACGAGTGGGCTGATCT 4860
```

```
Db 2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT 2700

Qy 2701 GTCTCGTTTTGCTTCACTTGTTACCTGTGCTTGTTAGTTAACTCCTATCT 2760
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db 2701 GTCTCGTTTTGCTTCACTTGTTACCTGTGCTTGTTAGTTAACTCCTATCT 2760

Qy 2761 TTGGCAACGTTGGGAACAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTTCCT 2820
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761 TTGGCAACGTTGGGAACAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTTCCT 2820

Qy 2821 TGTGCTGGTTGTTGTTCCCCGGTGCGACATATGACGCGCTGGTGACTTTCTGTGTGTCA 2880
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2821 TGTGCTGGTTGTTGTTCCCCGGTGCGACATATGACGCGCTGGTGACTTTCTGTGTGTCA 2880

Qy 2881 CGTAGCTCTTCTATGTTTAAACATCCAGTGCGCAGACTGTCTTGGGACTCTAGGGT 2940
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2881 CGTAGCTCTTCTATGTTTAAACATCCAGTGCGCAGACTGTCTTGGGACTCTAGGGT 2940

Qy 2941 TAGGGCCCATAGAATGTTGGTGCGTCTCCGGAAAGTGTCATGCTTGGTATTCTCATTATGT 3000
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2941 TAGGGCCCATAGAATGTTGGTGCGTCTCCGGAAAGTGTCATGCTTGGTATTCTCATTATGT 3000

Qy 3001 TCTTAAGTTTTTCCTCTTAGTGTTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA 3060
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3001 TCTTAAGTTTTTCCTCTTAGTGTTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA 3060

Qy 3061 TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCATTTTTCCC 3120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3061 TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCATTTTTCCC 3120

Qy 3121 TTTTGAAAGGCAAGGCAAGGGTCTATAGGAATGGAATGAAGGAAGACGCTTGGCGTGTGGGGACAC 3180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGGAATGAAGGAAGACGCTTGGCGTGTGGGGACAC 3180

Qy 3181 GGTTGATGGTTGCCCGTGTTGTTTTCGCGACCTTGTTTCGCAGTGTCTCTCTGAACGTGG 3240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3181 GGTTGATGGTTGCCCGTGTTGTTTTCGCGACCTTGTTTCGCAGTGTCTCTCTGAACGTGG 3240

Qy 3241 GCCGCCAGATGGTGGGCGCCATTACCGCCACCTTTACGCGTGCGAGTGTCTCTCTGACGTGG 3300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3241 GCCGCCAGATGGTGGGCGCCATTACCGCCACCTTTACGCGTGCGAGTGTCTCTCTGACGTGG 3300

Qy 3301 CACGCTGTCAGCGATGGCAGTGCCCGAACCCCGAACTTGGACTGGAAC 3360
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db 3301 CACGCTGTCAGCGATGGCAGTGCCCGAACCCCGAACTTGGACTGGAAC 3360

Qy 3361 TATCTTCAGATTAGGATCTCTGGCCACTAGTCACATGGGACTACATGGGGATTTGTTTGTGACAACGTGTT 3420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3361 TATCTTCAGATTAGGATCTCTGGCCACTAGTCACATGGGACTACATGGGGATTTGTTTGTGACAACGTGTT 3420

Qy 3421 GTATACTGCTCACCATGGCAGCAGGGGCGCCGCCGGTTGGCTCATCCCACCAGGCTCTATACA 3480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3421 GTATACTGCTCACCATGGCAGCAGGGGCGCCGCCGGTTGGCTCATCCCACCAGGCTCTATACA 3480

Qy 3481 CCCAATAACCGTTGACGCGGCTAATGCAGGCGGCTAATCTATCAACCACCATGTGAGCTGG 3540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3481 CCCAATAACCGTTGACGCGGCTAATGCAGGCGGCTAATCTATCAACCACCATGTGAGCTGG 3540

Qy 3541 GTCCCTTACTCGGTGCTCTTGCGGGAGACCAAGGGTATCGGGTAACACGACTGGGGTC 3600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3541 GTCCCTTACTCGGTGCTCTTGCGGGAGACCAAGGGTATCGGGTAACACGACTGGGGTC 3600

Qy 3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGGCGGGGCATGTTATTGG 3660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGGCGGGGCATGTTATTGG 3660

Qy 3661 GGCTGTTGCCAAGGGTTCTTCGCCAGGTTCTTCAGGTTCTTCAGGTTCTTCAGGTTCTTCAGGTTCTTATTGG 3720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3661 GGCTGTTGCCAAGGGTTCTTCGCCAGGTTCTTCAGGTTCTTCAGGTTCTTCAGGTTCTTCAGGTTCTTATTGG 3720

Qy 3721 GATGTTCACCCGTTCAGTCAGGTTAGGGTTAGGCCGTT 3780
           ||||||||||||||||||||||||||||||||||||||
Db 3721 GATGTTCACCCGTTCAGTCAGGTTAGGGTTAGGCCGTT 3780
```

```
Qy  4861 CTTTTCATGGTCAACCCCGAACCTTCATTTGTCAATACTGCAAAAAGAACTGCTGACAA 4920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861 CTTTTCATGGTCAACCCCGAACCTTCATTTGTCAATACTGCAAAAAGAACTGCTGACAA 4920

Qy  4921 TTATGTTTGTTGACTGAGCGCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCC 4980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921 TTATGTTTGTTGACTGAGCGCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCC 4980

Qy  4981 CAATGAGCGCACCACGGTGGCAGGGAGCCCGGCGTGGGAAAAACCTTGTGTGGGTTCTGTG 5040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981 CAATGAGCGCACCACGGTGGCAGGGAGCCCGGCGTGGGAAAAACCTTGTGTGGGTTCTGTG 5040

Qy  5041 GCGCTTGACGCGCTGACGCGCCTGTCCTGGCCCAGAGCCCAGCGAGGTGACCAGATACCA 5100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041 GCGCTTGACGCGCTGACGCGCCTGTCCTGGCCCAGAGCCCAGCGAGGTGACCAGATACCA 5100

Qy  5101 AATGTGCTTCACTGAAGTCAATATCTTCGGACAGACCCGCACTCGCTGTTGGCGTTGGAGT 5160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101 AATGTGCTTCACTGAAGTCAATATCTTCGGACAGACCCGCACTCGCTGTTGGCGTTGGAGT 5160

Qy  5161 GGCTATGGCTTATCTAGCCATTGACACTTTGGCGCCGCCACTTTGTGTGCGCGGTTGCTGTC 5220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161 GGCTATGGCTTATCTAGCCATTGACACTTTGGCGCCGCCACTTTGTGTGCGCGGTTGCTGTC 5220

Qy  5221 TATTACATCAGTCCCTACCGGTGCTACTGTCGTCGGGTTGACGAAGAAGAAATCGT 5280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221 TATTACATCAGTCCCTACCGGTGCTACTGTCGTCGGGTTGACGAAGAAGAAATCGT 5280

Qy  5281 GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGGTTGCTGCAATTGACAAGCTGAA 5340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281 GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGGTTGCTGCAATTGACAAGCTGAA 5340

Qy  5341 GAGTACAACATCACCACAGTCCTTTCACATTGAAAACCGCCCTTGAAAAAACTTAACAC 5400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341 GAGTACAACATCACCACAGTCCTTTCACATTGAAAACCGCCCTTGAAAAAACTTAACAC 5400

Qy  5401 CTTTCTTGGGCCTCATGGAGCGCTAGAGCGGCCACTGGGCGTTCATCAGAGTATTGCTGTGGTTTAGT 5460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401 CTTTCTTGGGCCTCATGGAGCGCTAGAGCGGCCACTGGGCGTTCATCAGAGTATTGCTGTGGTTTAGT 5460

Qy  5461 CACTTTACCTGGTTTGTCTTTGCATCATGCGTGTTTCATTGCGGGTATTACTAC 5520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461 CACTTTACCTGGTTTGTCTTTGCATCATGCGTGTTTCATTGCGGGTATTACTAC 5520

Qy  5521 CCCACTACCTCACAAGATCAAAGTGTTCCGTCATTATTGGAGGCCGCAATTGCGTCCAA 5580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521 CCCACTACCTCACAAGATCAAAGTGTTCCGTCATTATTGGAGGCCGCAATTGCGTCCAA 5580

Qy  5581 GCTTACAGACGCTAGAGGCGCACTGGGCGTTCAATCCGGCCGAGTTGTGGGCGTCTTGTCAGC 5640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581 GCTTACAGACGCTAGAGGCGCACTGGGCGTTCAATCCGGCCGAGTTGTGGGCGTCTTGTCAGC 5640

Qy  5641 TTGTACAATGGACATCGGTGGTTTGTCTTTGACAACAGCACTGGCGTATGCTGCCGCCTC 5700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641 TTGTACAATGGACATCGGTGGTTTGTCTTTGACAACAGCACTGGCGTATGCTGCCGCCTC 5700

Qy  5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGCCCACTATGGATCAGCT 5760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGCCCACTATGGATCAGCT 5760

Qy  5761 TGCTGGTTTAGTCTACTCCGCGTTCAATCCGGCCGAGTTGCTTCGCTTAGGCGTCTTGTCAGC 5820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761 TGCTGGTTTAGTCTACTCCGCGTTCAATCCGGCCGAGTTGCTTCGCTTAGGCGTCTTGTCAGC 5820

Qy  5821 AGGAGTGTCATGTCCTAGGAGCAACGACAGCAGGGCCAGATCACTGGCCAACAGACTTCTTAC 5880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821 AGGAGTGTCATGTCCTAGGAGCAACGACAGCAGGGCCAGATCACTGGCCAACAGACTTCTTAC 5880

Qy  5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTAACTTTATTGCCACTCGTTGACATCCG 5940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTAACTTTATTGCCACTCGTTGACATCCG 5940

Qy  5941 CAGGAAGATACTGGGCATTCTGGAGGCATTCTACCCCCTGGAGTGTCATATCAGCTTGCAT 6000
```

```
Qy  5941 CAGGAAGATACTGGGCATTCTGGAGGCATTCTACCCCCTGGAGTGTCATATCAGCTTGCAT 6000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941 CAGGAAGATACTGGGCATTCTGGAGGCATTCTACCCCCTGGAGTGTCATATCAGCTTGCAT 6000

Qy  6001 CCGTTGGCTCCACACCCGACGGAGGATGATTGCGGCGTCTCATTGGTGGGCTCTTAGAGAT 6060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001 CCGTTGGCTCCACACCCGACGGAGGATGATTGCGGCGTCTCATTGGTGGGCTCTTAGAGAT 6060

Qy  6061 TTGGCAGTATGTGCAATTCTTGTGTGATTGTTGCTTAAAGTCCTTAAAGCTGGAGTTCA 6120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061 TTGGCAGTATGTGCAATTCTTGTGTGATTGTTGCTTAAAGTCCTTAAAGCTGGAGTTCA 6120

Qy  6121 GAGCATGGTTAACATTCCTGGTGTCCTTTCTACAGCTGCCAGAGAGGGGTACAAGGGCCC 6180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121 GAGCATGGTTAACATTCCTGGTGTCCTTTCTACAGCTGCCAGAGAGGGGTACAAGGGCCC 6180

Qy  6181 CTGGAATTGGATCAGCTATGCTCCAAGCACGCGTGCTCCATGCGGTGCTGCAACTCATCTTTTC 6240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181 CTGGAATTGGATCAGCTATGCTCCAAGCACGCGTGCTCCATGCGGTGCTGCAACTCATCTTTTC 6240

Qy  6241 TGTTGAGAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG 6300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241 TGTTGAGAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG 6300

Qy  6301 AGGGGCTGTTCCAGTCAACGCTAGGCTGTGTGTCGGCTAGACCGTAGGCTGATTG 6360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301 AGGGGCTGTTCCAGTCAACGCTAGGCTGTGTGTCGGCTAGACCGTAGGCTGATTG 6360

Qy  6361 GACTAGTCTTGTCCTCAATTATGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA 6420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361 GACTAGTCTTGTCCTCAATTATGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA 6420

Qy  6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCAAC 6480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCCAAC 6480

Qy  6481 CTTGAGAGCTGCAGTGGCCGTGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481 CTTGAGAGCTGCAGTGGCCGTGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540

Qy  6541 AACTCCTTGGACGACATCTGCTTGCTCCTGACGGTAAGGGTAAGGTAAACTGTTAA 6600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541 AACTCCTTGGACGACATCTGCTTGCTCCTGACGGTAAGGGTAAGGTAAACTGTTAA 6600

Qy  6601 GCTTCCTTCCGCGTTCCGCGGTTGACGGTGTGCGGCCATGCGGCGCAACTTAATTTGCGTGA 6660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601 GCTTCCTTCCGCGTTCCGCGGTTGACGGTGTGCGGCCATGCGGCGCAACTTAATTTGCGTGA 6660

Qy  6661 TGCACTTGAGACAAATGACTGTAATTCCACAAACAACACTCCTAGTGATGAAGGCGCAGT 6720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661 TGCACTTGAGACAAATGACTGTAATTCCACAAACAACACTCCTAGTGATGAAGGCGCAGT 6720

Qy  6721 GTCCGCTCTTGTTTCAAACAGGAGTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC 6780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721 GTCCGCTCTTGTTTCAAACAGGAGTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC 6780

Qy  6781 AGCTGGCGGTTGACACCACCAAACTGCCAGCCCCTCATCGAAGAGTAGTGGTAAGAAA 6840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781 AGCTGGCGGTTGACACCACCAAACTGCCAGCCCCTCATCGAAGAGTAGTGGTAAGAAA 6840

Qy  6841 GCGCCAGTTCCGGGCAAGAAGAGTTGTGCCGTCTTGTCAGCGCCCGAGATCCGTCCC 6900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841 GCGCCAGTTCCGGGCAAGAAGAGTTGTGCCGTCTTGTCAGCGCCCGAGATCCGTCCC 6900

Qy  6901 AGGAGTGTCATGTCCTGAAAAGCTGCAACGAAGTGACCCGTTGAGAAGTCCTTCAAACCT 6960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901 AGGAGTGTCATGTCCTGAAAAGCTGCAACGAAGTGACCCGTTAGAAGTCCTTTAAACCT 6960

Qy  6961 CCCTCCTTCACCACTGTTCTACAGTTGCCATGCCGATGCCCGTTGGGAGCGGGGTGA 7020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961 CCCTCCTTCACCACTGTTCTACAGTTGCCATGCCGATGCCCGTTGGGAGCGGGGTGA 7020

Qy  7021 GTGTAACCCTTTCACTGGTAATGTGCAATGACCGGAAACAGGCGGAGCCCTGATGA 7080
```

Qy 8161 TGCAGCCGAACAGGCTTGGCATGAAGAGACCCTGCTTCCTTATTGCGGCGATGATTGCAC 8220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8161 TGCAGCCGAACAGGCTTGGCATGAAGAGACCCTGCTTCCTTATTGCGGCGATGATTGCAC 8220

Qy 8221 CGTAATTTGGAAGAGCGCCGGAGCAGAGTGCAGACAAACAAGCAATGGCGTGTCTTTGCTAG 8280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8221 CGTAATTTGGAAGAGCGCCGGAGCAGAGTGCAGACAAACAAGCAATGGCGTGTCTTTGCTAG 8280

Qy 8281 CTGGATGAAGGTGATGGGTGCACCACAAGATGTGTGCCTCAACCCAAATACAGTTTGGA 8340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8281 CTGGATGAAGGTGATGGGTGCACCACAAGATGTGTGCCTCAACCCAAATACAGTTTGGA 8340

Qy 8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400

Qy 8401 CTACTTCTTACAAGAGATCTCGTATCCCCCTGGCAGTGGTCTCTGCCGAGGTGCTCTGGG 8460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8401 CTACTTCTTACAAGAGATCTCGTATCCCCCTGGCAGTGGTCTCTGCCGAGGTGCTCTGGG 8460

Qy 8461 ATACAACCCAGTGCTGCGTGGTATCTAATACATCACTACCCATGTTTGTGGGT 8520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8461 ATACAACCCAGTGCTGCGTGGTATCTAATACATCACTACCCATGTTTGTGGGT 8520

Qy 8521 TAGCCGTGTGTTGGCGTCGTCCATTCATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCGA 8580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8521 TAGCCGTGTGTTGGCGTCGTCCATTCATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCGA 8580

Qy 8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCTAAGTC 8640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8581 GACGGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCTAAGTC 8640

Qy 8641 CATCATTGCTGGTGTGCACGGTATTGAGGCTTTCTCGGTGGTGCGCTACACCAACGCTGA 8700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8641 CATCATTGCTGGTGTGCACGGTATTGAGGCTTTCTCGGTGGTGCGCTACACCAACGCTGA 8700

Qy 8701 GATCCTCAGAGTTCCCAATCACTACTAACAGATGACCATGCCCCCTGCGAGCCTGGCG 8760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8701 GATCCTCAGAGTTCCCAATCACTACTAACAGATGACCATGCCCCCTGCGAGCCTGGCG 8760

Qy 8761 AAAGAAAGCCAGGCGGTCCTCGCCATGCGCAGCGCCAAGAGAACCAACAGGGCACGCCAAAATT 8820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8761 AAAGAAAGCCAGGCGGTCCTCGCCATGCGCAGCGCCAAGAGAACCAACAGGGCACGCCAAAATT 8820

Qy 8821 GGCTCGCTTCCTTCTCTCGGCATGCTACATCTCAGACCTTCTAGACTTGGATAAGACGAG 8880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8821 GGCTCGCTTCCTTCTCTCGGCATGCTACATCTCAGACCTTCTAGACTTGGATAAGACGAG 8880

Qy 8881 CGTGGCTCGGTACACCACCTTTCAATTATTGTGAAGTTCCTGTGAAGTATTGGCTGTCATTGTTT 8940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8881 CGTGGCTCGGTACACCACCTTTCAATTATTGTGAAGTTCCTGTGAAGTATTGGCTGTCATTGTTT 8940

Qy 8941 TATTACACCAGAAGATTGCAGAAGTTCTTGTGAAGTATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8941 TATTACACCAGAAGATTGCAGAAGTTCTTGTGAAGTATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9000

Qy 9001 TGCCCTAGGGCTCATTGCTGGATTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9001 TGCCCTAGGGCTCATTGCTGGATTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060

Qy 9061 CTAACAGTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGGCCAACAGGGGAGGGAGAGGGGC 9120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9061 CTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGGCCAACAGGGGAGGGAGACCCGGGC 9120

Qy 9121 TTAACGACCCGCCGATGTGAGTTTGGCGACCATGGTTGGCGACCATGGTTCGGGTGAA 9180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9121 TTAACGACCCGCCGATGTGAGTTTGGCGACCATGGTTGGCGACCATGGTTCGGGTGAA 9180

Qy 9181 GCCATGGTCTGAAGGGGATGACGTCCCTTCTCGGCTCATCCACAAAAACCGTCCGGGTGG 9240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9181 GCCATGGTCTGAAGGGGATGACGTCCCTTCTCGGCTCATCCACAAAAACCGTCCGGGTGG 9240

Db 7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAACAGGCGGAGGCCCTGATGA 7080

Qy 7140 TTTACCCAGTACCCTCCCAAAAAGAGGAGGTCTCTGAATGGTCAGACGAAAGTTGGTCGAC 7140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7081 TTTCCACTCAGTACCCTCCCAAAAAGAGGAGGTCTCTGAATGGTCAGACGAAAGTTGGTCGAC 7140

Qy 7200 GGCTACAACCGCTTCCAGCTACGTTACTGGCCCCCGTACCCTAAGATACGGGGAAAGGA 7200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7141 GGCTACAACCGCTTTCCAGCTACGTTACTGGCCCCCGTACCCTAAGATACGGGGAAAGGA 7200

Qy 7260 TTCCACTCAGTCAGCCCCGCGCCAAACGGCCTACAAAAAGAAGAAGTTGGGAAAGAGTGAGTT 7260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7201 TTCCACTCAGTCAGCCCCGCGCCAAACGGCCTACAAAAAGAAGAAGTTGGGAAAGAGTGAGTT 7260

Qy 7320 TTCGTGCAGCATGAGCTACACCTGGACCACGACTGATTGAGCTTCAAAAACTGCTTCTAAAGT 7320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7261 TTCGTGCAGCATGAGCTACACCTGGACCACGACTGATTGAGCTTCAAAAACTGCTTCTAAAGT 7320

Qy 7380 TCTGTCTGCAACTCGGGCCATCACTAGTGGTTTCCTCAAACAAGATCATTGGTGTATGT 7380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7321 TCTGTCTGCAACTCGGGCCATCACTAGTGGTTTCCTCAAACAAGATCATTGGTGTATGT 7380

Qy 7440 GACTGAGCCGCGGGATGCGGGAGCTTAGAGAAACAAGAAAAAGTCACTATTAATAGACAACCTCT 7440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7381 GACTGAGCCGCGGGATGCGGGAGCTTAGAGAAACAAGAAAAAGTCACTATTAATAGACAACCTCT 7440

Qy 7500 GTTCCCCCATCATACCACCACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7441 GTTCCCCCATCATACCACCACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500

Qy 7560 CGGTGTCATGTGGGACTATGATGAGCAGCTCGATGTCTGTTCGGGGCACTGATGTTCGTCT 7560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7501 CGGTGTCATGTGGGACTATGATGAGCAGCTCGATGTCTGTTCGGGGCACTGATGTTCGTCT 7560

Qy 7620 CCACATCACTGGCCTTCGGGGCACTTCGTTCGTGATGTTCGTTCTCTGGAGCAGCCCGAAGGCTGTTCT 7620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7561 CCACATCACTGGCCTTCGGGGCACTTCGTTCGTGATGTTCGTTCTCTGGAGCAGCCCGAAGGCTGTTCT 7620

Qy 7680 GGACTTGCGAGAGTGTGTCGAGGCAGTGGCAGTGTCAGAGTACCGGAGTCATTATCGGCAAACTGTGAT 7680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7621 GGACTTGCGAGAGTGTGTCGAGGCAGTGGCAGTGTCAGAGTACCGGAGTCATTATCGGCAAACTGTGAT 7680

Qy 7740 AGTTCCAAAGAGGAGAGGTCTTCGTGAAGACCCCCAGAGAAACCAACAAAGAAACCCCCAAG 7740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7681 AGTTCCAAAGAGGAGAGGTCTTCGTGAAGACCCCCAGAGAAACCAACAAAGAAACCCCCAAG 7740

Qy 7800 GCTTATCTCGTACCCCCACCTTGAAAGTGTCATGGGAGATGTGTTGAGAAGATGCCTACTACGGTCAGGT 7800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7741 GCTTATCTCGTACCCCCACCTTGAAAGTGTCATGGGAGATGTGTTGAGAAGATGCCTACTACGGTCAGGT 7800

Qy 7860 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTCATTGCGTGAGATCCACGTAC 7860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTCATTGCGTGAGATCCACGTAC 7860

Qy 7920 CCGTGTCAAGCGTCTGTTGTCGGATGTGTCACCCGGATGCGGAGCCACATGCGGATAC 7920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7861 CCGTGTCAAGCGTCTGTTGTCGGATGTGTCACCCGGATGCGGAGCCACATGCGGATAC 7920

Qy 7980 AGTGTGTTTGACAGTTACCATCACCACACCCGAGCTGGAGATATCATGGGAGACACAGCATCTACTC 7980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7921 AGTGTGTTTGACAGTTACCATCACCACACCCGAGCTGGAGATATCATGGGAGACACAGCATCTACTC 7980

Qy 8040 AGCAGCTAAACTCAGTGACCAACACCCGAGCTGGCATTCACACCATTGCGAGGCAGTATA 8040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7981 AGCAGCTAAACTCAGTGACCAACACCCGAGCTGGCATTCACACCATTGCGAGGCAGTATA 8040

Qy 8100 CGGCTGGAGGACCGATGATGTTGACCTCAAGTTCAACAGTTCAACACCTTTGACCTGCGCTGCTAAATGC 8100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8041 CGGCTGGAGGACCGATGATGTTGACCTCAAGTTCAACAGTTCAACACCTTTGACCTGCGCTGCTAAATGC 8100

Qy 8160 TTCCGGGGTCTATACTACCTCAAGTTCAACAGTTCAACACCTTTGACCTGCGCTGCTAAATGC 8160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8101 TTCCGGGGTCTATACTACCTCAAGTTCAACAGTTCAACACCTTTGACCTGCGCTGCTCATCCTGCGCTGCTAAATGC 8160

CC ribosome entry site, NS3-NS5B sequence, and GBV-B 3' UTR. Accordingly,
CC they are useful in providing tools for studying GBV-B replication,
CC polyprotein production and processing, identifying compounds that inhibit
CC GBV-B, providing a surrogate model for identifying compounds that inhibit
CC HCV, and providing a scaffold for producing GBV-B/HCV chimeric replicons.
CC Due to the similarity between GBV-B and the hepatitis C virus (HCV),
CC compounds that inhibit GBV-B may be useful antiviral agents, specifically
CC anti-HCV agents. The GBV-B subgenomic replicon constructs termed GBV-B-
CC neo-RepA (neo-RepA), neo-RepB, neo-RepC and neo-RepD were produced by
CC replacing the regions coding for structural proteins and the NS2 protein
CC with the sequences of neomycin phosphotransferase gene (neo) and
CC encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) in
CC the plasmid FL3/pACYC177. This polynucleotide sequence is the genomic GBV
CC -B replicon cDNA sequence of the invention.
XX
SQ Sequence 9397 BP; 2125 A; 2331 C; 2431 G; 2510 T; 0 U; 0 Other;

Query Match   99.4%;   Score 9338.6;   DB 9;   Length 9397;
Best Local Similarity 99.7%;   Pred. No. 0;
Matches 9368;   Conservative 0;   Mismatches 29;   Indels 2;   Gaps 1;

```
Qy       1 ACCACAAACACTCCAGTTGTTACACTCCGTAGGAATGCTCTGGAGCACCCCCCTAG 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ACCACAAACACTCCAGTTGTTACACTCCGTAGGAATGCTCTGGAGCACCCCCCTAG 60

Qy      61 CAGGGCGTGGGGGATTCCCCTGCCCGTCTGCGAGAAGGGTGGAGCCAACCACCTTAGTAT 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CAGGGCGTGGGGGATTCCCCTGCCCGTCTGCGAGAAGGGTGGAGCCAACCACCTTAGTAT 120

Qy     121 GTAGGCGGCGGGACTCATGAGCGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GTAGGCGGCGGGACTCATGAGCGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC 180

Qy     181 CCTGATGGGCGTTCATGGGTTCGTGGTGGTGGTGCGCGCTTAGGCAGCCTCCACGCCCACCA 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CCTGATGGGCGTTCATGGGTTCGTGGTGGTGGTGCGCGCTTAGGCAGCCTCCACGCCCACCA 240

Qy     241 CCTCCCAGATAGAGCGGGCACTGTAGGGAAGACCGGGAGACCGGTCACTACCAAGGACG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CCTCCCAGATAGAGCGGGCACTGTAGGGAAGACCGGGAGACCGGTCACTACCAAGGACG 300

Qy     301 CAGACCTCTTTTGAGTATCACGCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTGT 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CAGACCTCTTTTGAGTATCACGCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTGT 360

Qy     361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACCTACTGCCTGATAGGGTCCTTGCGAGGGGAT 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACCTACTGCCTGATAGGGTCCTTGCGAGGGGAT 420

Qy     421 CTGGGAGTCTCGTAGACCGTAGCACATGCCGTGTTATTCTACTCAAACAAGTCCTGTACC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CTGGGAGTCTCGTAGACCGTAGCACATGCCGTGTTATTCTACTCAAACAAGTCCTGTACC 480

Qy     481 TGCGCCCAGAACGCGCAAGAACAAGCAGAGCAGGCTTCATATCCGTGTCATTAAAAC 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGCGCCCAGAACGCGCAAGAACAAGCAGAGCAGGCTTCATATCCGTGTCATTAAAAC 540

Qy     541 ATCTGTTGAAAGGGGACAACGACGAGCAAAGCGCAAAGTCCAGCGCGATGCGTCGGCCTCGTAA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 ATCTGTTGAAAGGGGACAACGACGAGCAAAGCGCAAAGTCCAGCGCGATGCGTCGGCCTCGTAA 600

Qy     601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTTGCCAGC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTTGCCAGC 660

Qy     661 TCATGGTTGGGGACGCCAAGACCCTGGCCCATAAGTCTCGCCAATCTTGGAATCCTTCTGA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 TCATGGTTGGGGACGCCAAGACCCTGGCCCATAAGTCTCGCCAATCTTGGAATCCTTCTGA 720

Qy     721 TTACCCTTTGGGTGTGATTGGTTACAACTGTTACAACCACCACACCCTTCAGTAGGCCCGCTGGT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TTACCCTTTGGGTGTGATTGGTTACAACTGTTACAACCACCACACCCTTCAGTAGGCCCGCTGGT 780
```

```
Qy    9241 GTGAGGAGTCCTGGCTGTGTGGGAAGCAGTCAGTATAAATTCCCGTCGTGTCGTGACGC 9300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9241 GTGAGGAGTCCTGGCTGTGTGGGAAGCAGTCAGTATAAATTCCCGTCGTGTCGTGGTGACGC 9300

Qy    9301 CTCACGACGTATTTGTCCGCTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTG 9360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9301 CTCACGACGTATTTGTCCGCTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTG 9360

Qy    9361 TTCCAAGCGGAGGGCAACCCCCGCTTGGAATTAAAAACT 9399
           |||||||||||||||||||||||||||||||||||||||
Db    9361 TTCCAAGCGGAGGGCAACCCCCGCTTGGAATTAAAAACT 9399
```

RESULT 4
ADA77752
ID ADA77752 standard; cDNA; 9397 BP.
XX
AC ADA77752;
XX
DT 20-NOV-2003 (first entry)
XX
DE Genomic hepatitis GB virus B (GBV-B) replicon cDNA sequence.
XX
KW GB virus-B; GBV-B; replicon; autonomous replication; NS3-NS5B; HCV;
KW hepatitis C virus; antiviral; anti-HCV; neomycin phosphotransferase gene;
KW neo; ss.
XX
OS Hepatitis GB virus B.
XX
FH Key   Location/Qualifiers
FT 5'UTR   1..445
FT   /*tag= a
FT CDS   446..9040
FT   /*tag= b
FT   /product= "GBV-B polyprotein"
FT   /note= "From core protein to non structural protein 5B"
FT CDS   7268..9040
FT   /*tag= c
FT   /partial
FT   /product= "Non structural protein 5B (NS5B)"
FT   /note= "NS5B is an RNA dependent RNA polymerase; start
FT   codon is absent"
FT 3'UTR   9038..9397
FT   /*tag= d
XX
PN WO200305944-A2.
XX
PD 24-JUL-2003.
XX
PF 13-JAN-2003; 2003WO-EP000281.
XX
PR 15-JAN-2002; 2002US-0348573P.
PR 06-JUN-2002; 2002US-0386655P.
XX
PA (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
PI De Tomassi A, Graziani R, Paonessa G, Traboni C;
XX
DR WPI; 2003-598503/56.
DR P-PSDB; ADA77730.
XX
PT New GB virus B (GBV-B) replicon for identifying compounds that inhibit
PT GBV-B or hepatitis C virus, comprises a GBV-B 5' UTR, a selection or
PT reporter sequence, an internal ribosome entry site, an NS3-NS5B sequence,
PT and a GBV-B 3' UTR.
XX
PS Claim 3; Fig 2; 81pp; English.
XX
CC This invention relates to a novel GB virus-B (GBV-B) replicon and
CC replicon enhanced cells. A GBV-B replicon is an RNA molecule able to
CC autonomously replicate in a cultured cell to produce detectable levels of
CC one or more GBV-B proteins. Specifically, it may comprise the GBV-B 5'
CC UTR, GBV-B structural region, selection or reporter sequence, internal

```
Qy    781 GGCAGGAGCGGCTCGTTCGACCAGTCTGCCAGATAGTACGCTTGCTGGCGGATGAGTCAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GGCCAGGAGCCGGCTCGTTCGACCAGTCTGCCAGATAGTACGCTTGCTGGCGGATGAGTCAA 840

Qy    841 CTGGGCTACTGGTTGGTTCGGTGTCACCTTTTGTGTGGTATGTCTGCTATCTTTGGCCTG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTGGGCTACTGGTTGGTTCGGTGTCACCTTTTGTGTGGTATGTCTGCTATCTTTGGCCTG 900

Qy    901 TCCCTGTAGTGGGCGCGGGGTCACTGACCCAGACACAAATACCACCAATTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCCCTGTAGTGGGCGCGGGGTCACTGACCCAGACACAAATACCACCAATTG 960

Qy    961 CTGCCAGCGTAATCAGGTTATCTAGTTGTTCTCCTTCCACTTGCCTACACGAGCCTGGTTG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CTGCCAGCGTAATCAGGTTATCTAGTTGTTCTCCTTCCACTTGCCTACACGAGCCTGGTTG 1020

Qy   1021 TGTGATCTGTGCGGACGAGTGCTGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TGTGATCTGTGCGGACGAGTGCTGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA 1080

Qy   1081 TTGGACTGGCACGGACTCCTTCTTGGCTGACCACATTTGTTATGGGCGCTCTTGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TTGGACTGGCACGGACTCCTTCTTGGCTGACCACATTTGTTATGGGCGCTCTTGT 1140

Qy   1141 GACCTTGTGACGCCCTTGACATTGGTGAGTTGTGTGTGTATTAGTCGGTGACTG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTGTGACGCGCCCTTGACATTGGTGAGTTGTGTGTGTATTAGTCGGTGACTG 1200

Qy   1201 GCTTGTCAGGGACTGGCTTATTCACATAGACTGGTACTTGTTACCTGGA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GCTTGTCAGGGACTGGCTTATTCACATAGACTGGTACTTGTTACCTGGA 1260

Qy   1261 AGTGCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGATGGCCGGCAAGGT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGTGCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGATGGCCGGCAAGGT 1320

Qy   1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATTGCGACTAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATTGCGACTAT 1380

Qy   1381 GTTTAGCAGTGTACACTACCTGGCCGGTTGGGCGCTCTGATCTACTACTGGTCCCTCGGGGCAA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GTTTAGCAGTGTACACTACCTGGCCGGTTGGGCGCTCTGATCTACTACTGGTCCCTCGGGGCAA 1440

Qy   1441 GTGGTATCAGTTGCTCTCCTAGGCTTACATAGAAGCGACCTCGGAAACCCCAT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GTGGTATCAGTTGCTCTCCTAGGCTTACATAGAAGCGACCTCGGAAACCCCAT 1500

Qy   1501 CAGGGGATGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCTTTGATGATACCATGTCC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CAGGGGATGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCTTTGATGATACCATGTCC 1560

Qy   1561 TTGCCAACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAACACTTA 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TTGCCAACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAACACTTA 1620

Qy   1621 CAGGGCTATCACTCTAGAGTATAAACAACTCCATATCTTGGTACCCCTATACAAATCCCTGG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CAGGGCTGTCACTCTAGAGTATAAACAACTCCATATCTTGGTACCCCTATACAAATCCCTGG 1680

Qy   1681 TGCGAGGGGATGTATGGTTAAATTCAAAGAATTGCCGTATTCGCAA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGCGAGGGGATGTATGGTTAAATTCAAAGAATTGCCGTATTCGCAA 1740

Qy   1741 TGTGCCATCGTACTGACTATGGGCACTGATGTCAGTGGAACGACACTCGCAACACTTA 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TGTGCCATCGTACTGACTATGGGCACTGATGTCAGTGGAACGACACTCGCAACACTTA 1800

Qy   1801 CGAAGCATGCCGTGTAACACCATGCCATGCCATGTCAACACCGGCTCAGCCCTGAA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CGAAGCATGCCGTGTAACACCATGCCATGCCATGTCAACACCGGCTCAGCCCTGAA 1860

Qy   1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC 1920

Qy   1921 AGGCCATTGTATTTGAGGATCAGATACCCTATAGTTACTTTATGACCCTGTGAA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AGGCCATTGTATTTGAGGATCAGATACCCTATAGTTACTTTATGACCCTGTGAA 1980

Qy   1981 TTCCACTCTCCTACACCGGAGAGTGGGCTAGGTGCCCGGTACCCCACCTGGGTACG 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TTCCACTCTCCTACACCGGAGAGTGGGCTAGGTGCCCGGTACCCCACCTGGGTACG 2040

Qy   2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100

Qy   2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATAATCAGTCTTATATTCCGCACGGGGTGC 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATAATCAGTCTTATATTCCGCACGGGGTGC 2160

Qy   2161 TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGTGCTAATTCTGTTGGGGTTGTGTGG 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGTGCTAATTCTGTTGGGGTTGTGTGG 2220

Qy   2221 CAGCAAGTGATCTTATTTTAGCCTACTGTCCCTTGTTTTGGGGCGCTTC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CAGCAAGTGATCTTATTTTAGCCTACTGTCCCTTGTTTTGGGGCGCTTC 2280

Qy   2281 TGGTTACCCTTTGCGTCCTGTGCTCCATCCCAGTCGTATCTCCAAGCTGGCTGGGGATGT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 TGGTTACCCTTTGCGTCCTGTGCTCCATCCCAGTCGTATCTCCAAGCTGGCTGGGGATGT 2340

Qy   2341 TTTGTCTAAAGTCAAGTAGCTCCTTTGCTTTGATTTTCTTCATCTGTGTCATCTCCG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TTTGTCTAAAGTCAAGTAGCTCCTTTGCTTTGATTTTCTTCATCTGTGTCATCTCCG 2400

Qy   2401 CTGCAGGCTACGTTATGCGTCGCCCATGGCTGCGTGGGGCTTGCCCCT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 CTGCAGGCTACGTTATGCGTCGCCCATGGCTGCGTGGGGCTTGCCCCT 2460

Qy   2461 AACTTTCTTGTTGCAGCAGCTGCTGCCCACCAGATTATGACTGGTGGTGCGACTGCT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 AACTTTCTTGTTGCAGCAGCTGCTGCCCACCAGATTATGACTGGTGGTGCGACTGCT 2520

Qy   2521 AGTGGCAGGGTTAGTTTTGTGGGGCCGTAACCGCATAGCTCGCTTGT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 AGTGGCAGGGTTAGTTTTGTGGGGCCGTAACCGCATAGCTCGCTTGT 2580

Qy   2581 AGGTCCTGGCCTCTGGTAGGCGTTTAACCCTCTGGTTGGTTACGCCTGCTTCAGC 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 AGGTCCTGGCCTCTGGTAGGCGTTTAACCCTCTGGTTGGTTACGCCTGCTTCAGC 2640

Qy   2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGTTGTCAT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 TTTTGACACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGTTGTCAT 2700

Qy   2701 GTCTCGTTTTGGCTTCTTGCTCACTTGTACCTCGCTGTGTGCTTTAGTTAACTCCTATCT 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GTCTCGTTTTGGCTTCTTGCTCACTTGTACCTCGCTGTGTGCTTTAGTTAACTCCTATCT 2760

Qy   2761 TTTGGCAACGTTGGGAGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTTTCCT 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 TTTGGCAACGTTGGGAGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTTTCT 2820

Qy   2821 TGTGCTCGGTTGTTGTTCCCCGGTGCACATATGACGCGCTGGTGACTTTCTGTGTGTCA 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 TGTGCTCGGTTGTTGTTCCCCGGTGCACATATGACGCGCTGGTGACTTTCTGTGTGTCA 2880

Qy   2881 CGTAGCTCTTCTATGTTTAACATCCAGTGCAGCATCCGTCTTCTTTGGGACTGACTCTAGGGT 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 CGTAGCTCTTCTATGTTTAACATCCAGTGCAGCATCCGTCTTCTTTGGGACTGACTCTAGGGT 2940

Qy   2941 TAGGGCCCATGAAATGTTTCCGGAAAGTGTCGCATGCTTGGTTGGTATTCCATTATGT 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  2941 TAGGGGCCATAGAATGTGTGGCGTCTCGGAAAGTGCCATGCTGGTATTCTCATTATTGT 3000
Qy  3001 TCTTAAGTTTTTCCTCTCTTAGTGTTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 TCTTAAGTTTTTCCTCTCTTAGTGTTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA 3060
Qy  3061 TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCAAGAGCCATTTTTTCCC 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCAAGAGCCATTTTTTCCC 3120
Qy  3121 TTTTGAAGGCAAGGCAAGGTCTATAGGAATGAAGGAAGACGCTTGGCGTGTGGGGACAC 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 TTTTGAAGGCAAGGCAAGGTCTATAGGAATGAAGGAAGACGCTTGGCGTGTGGGGACAC 3180
Qy  3181 GGTTGATGGTTTGCCCGTGTTGCGCGTCTCGGCCGACCTTGTTTGTGCCAGGGTTGGCTAT 3240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 GGTTGATGGTTTGCCCGTGTTGCGCGTCTCGGCCGACCTTGTTTGTGCCAGGGTTAGCTAT 3240
Qy  3241 GCCGCCAGATGGGTGGGCCATTACCGCACCTTTTACGCTGCAGTGTCTCTCTGAACGTGG 3300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 GCCGCCAGATGGGTGGGCCATTACCGCACCTTTTACGCTGCAGTGTCTCTCTGAACGTGG 3300
Qy  3301 CACGCTGTCAGCGGATGGCAGTGGTCATGACTCGTGTATAGACCCCGACTTGGACTGGAAC 3360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 CACGCTGTCAGCGGATGGCAGTGGTCATGACTCGTGTATAGACCCCGACTTGGACTGGAAC 3360
Qy  3361 TATCTTCAGATTAGGATCTCTGCCACTAGCTACATGGGATTTGTTTGTGACAACGTGTT 3420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 TATCTTCAGATTAGGATCTCTGCCACTAGCTACATGGGATTTGTTTGTGACAACGTGTT 3420
Qy  3421 GTATACTGCTCACCATGGCAGCAAGGGGTATCGGCCGTTGGCTCATCCCACAGGCTCTATACA 3480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421 GTATACTGCTCACCATGGCAGCAAGGGGTATCGGCCGTTGGCTCATCCCACAGGCTCCATACA 3480
Qy  3481 CCCAATAACCGTTGACGCGCTAATGACCAGGACATCTATCAACAACATGTGGAGCTGG 3540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481 CCCAATAACCGTTGACGCGCTAATGACCAGGACATCTATCAACAACATGTGGAGCTGG 3540
Qy  3541 GTCCCTTACTCGGTGCTCTTCGCGGCCAAGGGGTATCTGGTAACACGACTGGGGGTC 3600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541 GTCCCTTACTCGGTGCTCTTCGCGGCCAAGGGGTATCTGGTAACACGACTGGGGGTC 3600
Qy  3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGTGCGGGGCCCTTCCCAT 3660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGTGCGGGGCCCTTCCCAT 3660
Qy  3661 GGCTGTTGCCAAGGGTTCTTCAGGACGATTCTGTCGCTCCTCCGGGCATGTTATTGG 3720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661 GGCTGTTGCCAAGGGTTCTTCAGGACGATTCTGTCGCTCCTCCGGGCATGTTATTGG 3720
Qy  3721 GATGTTCACCGCGCTAGAAAATTCTGGACGGTTCAGTCAGTCCAGATTAGGGTTAGGCCGTT 3780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721 GATGTTCACCGCGCTAGAAAATTCTGGACGGTTCAGTCAGTCCAGATTAGGGTTAGGCCGTT 3780
Qy  3781 GGTGTGTGCTGGATACCATCCCGATACAGCACATGCCACTCTTGATACAAAAACCTAC 3840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781 GGTGTGTGCTGGATACCATCCCGATACAGCACATGCCACTCTTGATACAAAAACCTAC 3840
Qy  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATGCCCCCACTGGCAGCGGCAAGTCAAC 3900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 TGTGTTTGCCTAACGAGTATTCAGTGCAAATTTAATGCCCCCACTGGCAGCGGCAAGTCAAC 3900
Qy  3901 CAAATTACCACTTTCTTACACATGGCCAAATGCCAAGATGAGGTCTTCGGTCCTAAATCCCAGTGT 3960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 CAAATTACCACTTTCTTACACATGGCCAAATGCCTAAATCCCTTGGTCTCAAATCCCAGTGT 3960
Qy  3961 GGCTACAACAGCATCAATGCCAAAGTGCACGACGGGTACGCGGTGAATCGAAATTG 4020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 GGCTACAACAGCATCAATGCCAAAGTGCACGACGGGTACGCGGTGAATCGAAATTG 4020
Qy  4021 CTATTTTAATGGCAAATGTACCAACAACAGGGGGCTTCACTTACGTACACGCACATATGGCAT 4080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 CTATTTTAATGGCAAATGTACCAACAACAGGGGGCTTCACTTACGTACACGCACATATGGCAT 4080
```

```
Qy  4081 GTACCTGACCGGAGCATGTTCCCGGAACTATCATTGTGACGACACTTGCA 4140
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 GTACCTGACCGGAGCATGTTCCCGGAACTATCATTGTGACGACACTTGCA 4140
Qy  4141 TACCGATGCAACCACCGGTGTTGGGCATTGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA 4200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 TACCGATGCAACCACCGGTGTTGGGCATTGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA 4200
Qy  4201 TGTTTAGGTCAGTGGTTCTTGCCACGGCTACCCCCCCTGGAGTAATCCTACCACCACATGC 4260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201 TGTTTAGGTCAGTGGTTCTTGCCACGGCTACCCCCCCTGGAGTAATCCTACCACCACATGC 4260
Qy  4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAAAGAT 4320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAAAGAT 4320
Qy  4321 TAAGGAGAAAAATCTGAGAAAAGGGAGACACCTTATCTTGAGGCTACCAAAAAACACTG 4380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321 TAAGGAGAAAAATCTGAGAAAAGGGAGACACCTTATCTTGAGGCTACCAAAAAACACTG 4380
Qy  4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440
Qy  4441 ATGTGACATCTCAAAAATCCCTGAGGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441 ATGTGACATCTCAAAAATCCCTGAGGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500
Qy  4501 TACAGGGTACACTGGTGACTTTGATTCGTGTGATGACTGCTCATGTGTGATGAGAGGCAC 4560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501 TACAGGGTACACTGGTGACTTTGATTCGTGTGATGACTGCTCATGTGTGATGAGAGGCAC 4560
Qy  4561 CGACGCAGCCAAGGCATGGTATGGTTGTCATCAACAGAGCTCAAACTATTCTCGGACAC 4620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561 ATGCCATGTGACCTTGACCTTGACCCTACTTTCGTGTTCGTGTGCGGGTTTCAGC 4620
Qy  4621 AATAGTTAAAGGCCAGGCAGCCGCACAGCCGTGGGAGAGCGTGGCATATACTACTA 4680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621 AATAGTTAAAGGCCAGGCAGCCGCACAGCCGTGGGAGAGCGTGGCATATACTACTA 4680
Qy  4681 TGTAGACGGGAGTTGTACCCCTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681 TGTAGACGGGAGTTGTACCCCTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740
Qy  4741 CGACGCAGCCAAGGCATGGTATGGTTGTCATCAACAGAGCTCAAACTATTCTGGACAC 4800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741 CGACGCAGCCAAGGCATGGTATGGTTGTCATCAACAGAGCTCAAACTATTCTGGACAC 4800
Qy  4801 CTATCGCACCACCCAACCTGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT 4860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801 CTATCGCACCACCCAACCTGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT 4860
Qy  4861 CTTTTCTATGTGGTCAACCCCGAACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA 4920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861 CTTTTCAATGTGGTCAACCCCGAACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA 4920
Qy  4921 TTATGTTTTGTTGACTGCAGGCATGGTATGGTTCATCAGTATCGTCATACAACTGTCATCAGTATGCTGCTCC 4980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921 TTATGTTTTGTTGACTGCAGGCATGGTATGGTTCATCAGTATCGTCATACAACTGTCATCAGTATGCTGCTCC 4980
Qy  4981 CAATGACGGCACCACGTGCAGGGCGCTGACGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG 5040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981 CAATGACGGCACCACGTGCAGGGCGCTGACGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG 5040
Qy  5041 GCGCTTGGACGGGGCTGACGCGCTGTCCTGCGCCAGAGCCCAGAGGGTGACCAGATACCA 5100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041 GCGCTTGGACGGGGCTGACGCGCTGTCCTGCGCCAGAGCCCAGAGGGTGACCAGATACCA 5100
Qy  5101 AATGTGCTTCACTGAAGTCAATAACTTTCTGGGACAGCAGCCCGACTCGCTGCGCTGTTG 5160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101 AATGTGCTTCACTGAAGTCAATAACTTTCTGGGACAGCAGCCCGACTCGCTGCGCTGTTG 5160
```

```
Qy  5161  GGCTATGGCTTATCTAGCCATTGACACTTTGGCGCCACTTGTGTCGGCGGTC  5220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161  GGCTATGGCTTATCTAGCCATTGACACTTTGGCGCCACTTGTGTCGGCGGTC  5220

Qy  5221  TATTACATCAGTCCCTACACGGTGCTACTGTCGCCCAGTGGTTGACGAAGAAATCGT  5280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221  TATTACATCAGTCCCTACACGGTGCTACTGTCGCCCAGTGGTTGACGAAGAAATCGT  5280

Qy  5281  GGAGGAGTGTGCATCATTCATTCCCTGGAGCGCCATGGCCATTGCTCGCAATTGACAAGCTGAA  5340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281  GGAGGAGTGTGCATCATTCATTCCCTGGAGCGCCATGGCCATTGCTCGCAATTGACAAGCTGAA  5340

Qy  5341  GAGTACAATCACCACAACTAGTCCTTTCACATTGGAAACCGCCCTTGAAAAACTTAACAC  5400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341  GAGTACAATCACCACAACTAGTCCTTTCACATTGGAAACCGCCCTTGAAAAACTTAACAC  5400

Qy  5401  CTTTCTTGGGCCTCATGCAGCTACAACATCCTTGCTATCATGAGAGTATTGCTGTGGTTAGT  5460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401  CTTTCTTGGGCCTCATGCAGCTACAACATCCTTGCTATCATGAGAGTATTGCTGTGGTTAGT  5460

Qy  5461  CACTTACCTGACAAATCCCTTTGCATCATGCCGTGTTTCATTGCGGTATTACTAC  5520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461  CACTTACCTGACAAATCCCTTTGCATCATGCCGTGTTTCATTGCGGTATTACTAC  5520

Qy  5521  CCCACTACCTCACAAGATCAAAATGTCCTGTCATTATTGGAGGCGCAATTGCGTCCAA  5580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521  CCCACTACCTCACAAGATCAAAATGTCCTGTCATTATTGGAGGCGCAATTGCGTCCAA  5580

Qy  5581  GCTTACAGACGCTAGAGGCGCACTGGCGTTCAATCGGCCGGCGGCTGCGGCGCTGCGCCTCT  5640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581  GCTTACAGACGCTAGAGGCGCACTGGCGTTCAATCGGCCGGCGGCTGCGGCGCTGCGCCTCT  5640

Qy  5641  TGGTACATGGACATCGGTGGGTTTTGCTTTTGTCTTTGACATGCTAGGCGGCTATGCGCCGCCTC  5700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641  TGGTACATGGACATCGGTGGGTTTTGCTTTTGTCTTTGACATGCTAGGCGGCTATGCGCCGCCTC  5700

Qy  5701  ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGTGAGTGGCCACTATGGATCAGCT  5760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701  ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGTGAGTGGCCACTATGGATCAGCT  5760

Qy  5761  TGCTGGTTTAGTCTACTCCGGCGTTCAATCGGCCGGCGGCGAGTTGTGTGGCGTCTTGTCAGC  5820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761  TGCTGGTTTAGTCTACTCCGGCGTTCAATCGGCCGGCGGCGAGTTGTGTGGCGTTTGTCAGC  5820

Qy  5821  TTGTGCAATGTTGCTTTGACAACAGCAGGGCCAGATCACTGGCCCAACAGACTTCTAC  5880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821  TTGTGCAATGTTGCTTTGACAACAGCAGGGCCAGATCACTGGCCCAACAGACTTCTAC  5880

Qy  5881  TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCCACTCGTGACATCCG  5940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881  TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCCACTCGTGACATCCG  5940

Qy  5941  CAGGAAGATACTGGGCATTCGGAGCATCAGCGGGGTAACACTCACCCCCGAAAGGGGTACAAGGGCCC  6000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941  CAGGAAGATACTGGGCATTCGGAGCATCAGCGGGGTAACACTCACCCCCGAAAGGGGTACAAGGGCCC  6000

Qy  6001  CCGTTGGCTCCAACACCCCAGACGGAGGATGATTGCGGCGCCTCATTGGGGTCTAGAGAT  6060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001  CCGTTGGCTCCAACACCCCAGACGGAGGATGATTGCGGCGCCTCATTGGGGTCTAGAGAT  6060

Qy  6061  TTGGCAGTATGCGTGCAATTTCTTTGTGATTTGCTTTGACTTTAAATGTCCTTAAAGCTGGAGTTCA  6120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061  TTGGCAGTATGCGTGCAATTTCTTTGTGATTTGCTTTGACTTTAAATGTCCTTAAAGCTGGAGTTCA  6120

Qy  6121  GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAAGAGGGGTACAAGGGCCC  6180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121  GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAAGAGGGGTACAAGGGCCC  6180

Qy  6181  CTGGATTGGATCAGGTATGCTTCCAAGCACGCTGTCCATGCCGTGTCGAACTCATCTTTC  6240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181  CTGGATTGGATCAGGTATGCTTCCAAGCACGCTGTCCATGCCGTGTCGAACTCATCTTTC  6240

Qy  6241  TGTTGAGAATGGTTTTGCAAAACTTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241  TGTTGAGAATGGTTTTGCAAAACTTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300

Qy  6301  AGGGGCTGTTCCAGTCAAGCGTCGTGTGTGGTCGGCTAGCAGCCGCTAGAAATGGGAGA  6360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301  AGGGGCTGTTCCAGTCAAGCGTCGTGTGTGGTCGGCTAGCAGCCGCTAGAAATGGGAGA  6360

Qy  6361  GACTAGTCTTGTCGTCAATTATGGCGTTAGGACTACTGTAAATATGGAAAATGTGTTAA  6420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361  GACTAGTCTTGTCGTCAATTATGGCGTTAGGACTACTGTAAATATGGAAAATGTGTTAA  6420

Qy  6421  TCACATTTTGTTACAGCAGTATCCTCTCCAAAATGTCGTTTCACCCAGGTGCCCCAAC  6480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421  TCACATCTTTGTTACAGCAGTATCCTCTCCAAAATGTCGTTTCACCCAGGTGCCCCAAC  6480

Qy  6481  CTTGAGAGCTGCAGTGGCCGTGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA  6540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481  CTTGAGAGCTGCAGTGGCCGTGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA  6540

Qy  6541  AACTCCTGGACGACATCTGCTTCGTCGTTCCTGACGGTAAGGGTAAAACTGTTAA  6600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541  AACTCCTGGACGACATCTGCTTCGTCGTTCCTGACGGTAAGGGTAAAACTGTTAA  6600

Qy  6601  GCTTCCCTCCGGTTGACGGTCACACACTGTGTGCGCATGCAACTAATTTGCGTGA  6660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601  GTTCCCTCCGGTTGACGGTCACACACTGTGTGCGCATGCAACTAATTTGCGTGA  6660

Qy  6661  TGCACTTGAGACAAATGACTGTAATTCCACAAACACTCCAGTGATGAGCCGCAGT  6720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661  TGCACTTGAGACAAATGACTGTAATTCCACAAACACTCCAGTGATGAGCCGCAGT  6720

Qy  6721  GTCCGCTCTGTTGTTTCAAACAGGAGTTGGCGGCGTACAAACCAATTGCTTGAGGCAATTTC  6780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721  GTCCGCTCTGTTGTTTCAAACAGGAGTTGGCGGCGTACAAACCAATTGCTTGAGGCAATTTC  6780

Qy  6781  AGCTGGCGTTGACACCACCAAACTGCCAGCCCCTCCATCGAAGAGGTAGTGTAAGAAA  6840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781  AGCTGGCGTTGACACCACCAAACTGCCAGCCCCTCCATCGAAGAGGTAGTGTAAGAAA  6840

Qy  6841  GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCCCCTCCGATCCGTCCC  6900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841  GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCCCCTCCGATCCGTCCC  6900

Qy  6901  AGGAGTGTCATGTCCTGCAAGCTGCGACGAGTGACCCGGTTAGAAGGTCCTTCAAACCT  6960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901  AGGAGTGTCATGTCCTGCAAGCTGCGACGAGTGACCCGGTTAGAAGGTCCTTCAAACCT  6960

Qy  6961  CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGCCGTGTTGGGACGCGGGGTGA  7020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961  CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGCCGTGTTGGGAGCGCGGGGTGA  7020

Qy  7021  GTGTAACCCTTTCACTGCAATTGGGATGTGCAATGACCGGCGAAACAGGCGGAGGGCCCTGATGA  7080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021  GTGTAACCCTTTCACTGCAATTGGGATGTGCAATGACCGGCGAAACAGGCGGAGGGCCCTGATGA  7080

Qy  7081  TTTACCCAGTACCCTCCCAAAAAGGAGGAGTCTCTGAATGGTCAGACGTGGTCGAC  7140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081  TTTACCCAGTACCCTCCCAAAAAGGAGGAGTCTCTGAATGGTCAGACGTGGTCAAC  7140

Qy  7141  GGCTACAACCGCTCAGTGACGCTTACTGGCCCCAGCTAGGTAGTTGGAAAGAAGGA  7200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141  GGCTACAACCGCTCAGTGACGCTTACTGGCCCCAGCTAGGTAGTTGGGAAAGAAGGA  7200

Qy  7201  TTCCACTTCAGTCAGCAGCCCCGCCAAACGGCTACAACAGCGGCCAAACGGCCTACAAGGGCCC  7260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201  TTCCACTTCAGTCAGCAGCCCCGCCAAACGGCTACAACAGCGGCCAAACGGCCTACAAGGGCCC  7260

Qy  7261  TTCGTGCAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAAGT  7320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261  TTCGTGCAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAAGT  7320

Qy  7321  TCTGTCTGCAACTCGGGCATCACTAGTGGTTTCCTCAACAAAGATCATTGGTGTATGT  7380
```

Qy 8461 ATACAAACCCAGTGCTGCCGTGGGTGGGTATCCATAATAACATCACTACCCATGTTGTGGGT 8520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8461 ATACAAACCCAGTGCTGCCGTGGGTGGGTATCCATAATACATCACTACCCATGTTGTGGGT 8520

Qy 8521 TAGCCGTGTGTTGGCTGTCCATTCATGTGGTATCGGAGCAGATGCTCTTGAGGACAAACTTCCCGA 8580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8521 TAGCCGTGTGTTGGCTGTCCATTCATGTGGTATCGGAGCAGATGCTCTTGAGGACAAACTTCCCGA 8580

Qy 8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCGCCAG 8640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCGCCAG 8640

Qy 8641 CATCATTGCTGGTGTGCACGGTATTGAGGCTTTCTCGGTGCGCTACACCAACGCTGA 8700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8641 CATCATTGCTGGTGTGCACGGTATTGAGGCTTTCTCGGTGCGCTACACCAACGCTGA 8700

Qy 8701 GATCCTCAGAGTTCCCAATCACTCAACAGACATGACCATGCCCCCCTGCGAGCCTGGCG 8760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8701 GATCCTCAGAGTTCCCAATCACTCAACAGACATGACCATGCCCCCCTGCGAGCCTGGCG 8760

Qy 8761 AAAGAAAGCCAGGCGCGGTCTCGCCAGCGCCAAGAGGGCGGTGCCGAGCACGCGCAAAATT 8820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8761 AAAGAAAGCCAGGCGCGGTCTCGCCAGCGCCAAGAGGGCGGTGCCGAGCACGCGCAAAATT 8820

Qy 8821 GGGTCGCTTCCTTCTCTCGCATGCTACATCTAGACCTCTACCAGATTTGGATAAGACGAG 8880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8821 GGGTCGCTTCCTTCTCTCGCATGCTACATCTAGACCTCTACCAGATTTGGATAAGACGAG 8880

Qy 8881 CGTTGGCTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCGGAGGGGGATGTGTT 8940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8881 CGTTGGCTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCGGAGGGGGATGTGTT 8940

Qy 8941 TATTACACCACAGAGAAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTGTCATTGTTTT 9000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8941 TGTTACACCACAGAGAAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTGTCATTGTTTT 9000

Qy 9001 TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9001 TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060

Qy 9061 CTAACAGTTTTTTTTTTTTTTTTTTT-||TTTTTTTTTTTTTTTAGGGCAGCCGCAACAGGGGACCCCGGGC 9120
          ||||||||| |||||||||||||||||||||||||||||||||||||||||
Db 9061 CTAACAG--TTTTTTTTTTTTTTTTAGGGCAGCCGCAACAGGGGACCCCGGGC 9118

Qy 9121 TTAACGACCCCGCCGATGTGAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAA 9180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9119 TTAACGACCCCGCCGATGTGAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAA 9178

Qy 9181 GCCATGGTCTGAAGGGGATGACGTCCCTTCTGGCTCATCCACAAAAACCGTCTCGGGTGG 9240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9179 GCCATGGTCTGAAGGGGATGACGTCCCTTCTGGCTCATCCACAAAAACCGTCTCGGGTGG 9238

Qy 9241 GTGAGGAGTCCTGCTGTGTGGGAAGCAGTCAGTATAATTCCCGTCGTGTGGTGTGACGC 9300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9239 GTGAGGAGTCCTGCTGTGTGGGAAGCAGTCAGTATAATTCCCGTCGTGTGGTGTGACGC 9298

Qy 9301 CTCACGACGTATTTGTCCCGCTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTG 9360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 9299 CTCACGACGTACTTGTCCCGCTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTG 9358

Qy 9361 TTCCAAGCGGAGGGCAACCCCGCTTGGAATTAAAAACT 9399
          |||||||||||||||||||||||||||||||||||||
Db 9359 TTCCAAGCGGAGGGCAACCCCGCTTGGAATTAAAAACT 9397

RESULT 5
AAF23485
ID   AAF23485 standard; DNA; 9139 BP.
XX
AC   AAF23485;
XX
DT   11-SEP-2003 (revised)

Db 7321 TCTGTCTGCAACTCGGGCCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGT 7380

Qy 7381 GACTGAGCCGCGGGATGCGGGGATCCGGGACTTAGAAACAAAAAGTCACTATTAATAGACAACCTCT 7440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7381 GACTGAGCCGCGGGATGCGGGGATCCGGGACTTAGAAACAAAAAGTCACTATTAATAGACAACCTCT 7440

Qy 7441 GTTCCCCATCATACCACCACAAGCAAGTCAAGTTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7441 GTTCCCCATCATACCACCACAAGCAAGTCAAGTTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500

Qy 7501 CGGTGTCATGTGGGACTATGAAGTAGCAGCTCACACGCCCTCTCAAGTCTGCTAAGTC 7560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7501 CGGTGTCATGTGGGACTATGAAGTAGCAGCTCACACGCCCTCTCAAGTCTGCTAAGTC 7560

Qy 7561 CCACATCACTGCCTTCGGGGCACTGATGTTCGTTCTTCTGGAGCAGCCCGCAAGGCTGTTCT 7620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7561 CCACATCACTGCCTTCGGGGCACTGATGTTCGTTCTTCTGGAGCAGCCCGCAAGGCTGTTCT 7620

Qy 7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGGTCATTATCGGCAAACTGTGAT 7680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGGTCATTATCGGCAAACAGTGAT 7680

Qy 7681 AGTTCCAAAGGACAGGAGGTCTTCGTGAAGACACCCCCAGAAAACCAACAAAGAAACCCCCAAG 7740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7681 AGTTCCAAAGGACAGGAGGTCTTCGTGAAGACACCCCCAGAAAACCAACAAAGAAACCCCCAAG 7740

Qy 7741 GCTTATCGTCGTACCCCCCCTTGAAATGAGATGTGTTGAGAGATGTACTACGGTCAGGT 7800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7741 GCTCATCTCGTACCCCCCCTTGAAATGAGATGTGTTGAGAGATGTACTACGGTCAGGT 7800

Qy 7801 TGCTCCTGAGCGTAGTGAAAGCTGTGATGGGAGATGGCTACGGGTTTGTAGATCCACGTAC 7860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7801 TGCTCATCTCGAGCGTAGTGAAAGCTGTGATGGGAGATGGCTACGGGTTTGTAGATCCACGTAC 7860

Qy 7861 CCGTGTCAAGCCTCGTTGTCGATGTGTCAGTCGGACGCCCACATGCGCGATAC 7920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7861 CCGTGTCAAGCCTCGTTGTCGATGTGTCAGTCGGACGCCCACATGCGCGATAC 7920

Qy 7921 AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTGGAGACAGCAGCATCTACTC 7980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7921 AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTGGAGACAGCAGCATCTACTC 7980

Qy 7981 AGCAGCTAAACTCAGTGACCAACCACAGCTGCATTCACACCACCATTGCGAGGCAGTTATA 8040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7981 AGCAGCTAAACTCAGTGACCAACCACAGCTGCATTCACACCACCATTGCGAGGCAGTTATA 8040

Qy 8041 CGCTGGAGGACCAGATGCTTATGATGGCCAGAGATCGGATATCGTAGTGTGTAGGTC 8100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8041 CGCTGGAGGACCAGATGCTTATGATGGCCAGAGATCGGATATCGTAGTGTGTAGGTC 8100

Qy 8101 TTCCGGCGGCTCTATACTACCTCAAGTTCAACAGTTTGACCTGCTCGCTGGAAGGTAAATGC 8160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8101 TTCCGGCGGCTCTATACTACCTCAAGTTCAACAGTTTGACCTGCTCGCTGGAAGGTAAATGC 8160

Qy 8161 TGCAGCCGAACAGGCTGCATGGAAGAACCCTCGCTTCCTTATTTGCGGCGATGATTGCAC 8220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8161 TGCAGCCGAACAGGCTGCATGGAAGAACCCTCGCTTCCTTATTTGCGGCGATGATTGCAC 8220

Qy 8221 CGTAATTTGGAAGACCCGGAGAGCAGATGCAGACAAAGCAATGCCGTGTCTTTGCTAG 8280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8221 CGTAATTTGGAAAGACCCGGAGAGCAGATGCAGACAAAGCAATGCCGTGTCTTTGCTAG 8280

Qy 8281 CTGGATGAAGGTGATGGGTGCACCACAAGATTGTGTGCCTCACCCAAATACAGTTTTGGA 8340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8281 CTGGATGAAGGTGATGGGTGCACCACAAGATTGTGTGCCTCACCCAAATACAGTTTTGGA 8340

Qy 8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCACCAAAAGTGGCAAGCCTTA 8400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCACCAAAAGTGGCAAGCCTTA 8400

Qy 8401 CTACTTTCTTACAAGAGATCCTCCGTATCCCCCCTTGGCAGGTGCTCTGCCGAGGGTCTGGG 8460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8401 CTACTTTCTTACAAGAGATCCTCCGTATCCCCCCTTGGCAGGTGCTCTGCCGAGGGTCTGGG 8460

DT 21-MAR-2001 (first entry)
XX
DE GBV-B virus genome #2.
XX
KW GBV-B; hepatitis C virus; HCV; vaccine; ds.
XX
OS Hepatitis GB virus B.
XX
PN WO200075337-A1.
XX
PD 14-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-US015293.
XX
PR 04-JUN-1999; 99US-0137694P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Bukh J, Yanagi M, Emerson SU, Purcell RH;
XX
PR WPI; 2001-091214/10.
XX
DR Example: Page 75-78; 96pp; English.
XX
CC New infectious nucleic acids of the GB virus-B clone, useful for
PT indirectly studying the molecular properties of hepatitis C virus (HCV)
PT and in developing vaccines and therapeutics for HCV.
XX
PS Example: Page 75-78; 96pp; English.
XX
CC The present invention relates to GB virus-B. The nucleic acid molecules
CC of the invention are useful for indirectly studying the molecular
CC properties of hepatitis C virus (HCV). The infectious nucleic acid
CC sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used
CC in the development of vaccines and therapeutics for HCV. (Updated on 11-
CC SEP-2003 to standardise OS field)
XX
SQ Sequence 9139 BP; 2073 A; 2261 C; 2349 G; 2456 T; 0 U; 0 Other;

Query Match   97.2%; Score 9131.4; DB 4; Length 9139;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9132; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   1 ACCAAAACACTCCAGTTGTTACACTCCGCTAGGAATGCCTCCTGAGCACCCCCCTAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCAAAACAACACTCCAGTTGTTACACTCCGCTAGGAATGCCTCCTGAGCACCCCCCTAG 60

Qy   61 CAGGGCGTGGGGGATTCCCCTGCCCGTCTCGCAGAAGGGGTGGAGCCAACCACCTTAGTAT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CAGGGCGTGGGGGATTCCCCTGCCCGTCTCGCAGAAGGGGTGGAGCCAACCACCTTAGTAT 120

Qy   121 GTAGGCGGCGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGGTTGACTTGATGGC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTAGGCGGCGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGGTTGACTTGATGGC 180

Qy   181 CCTGATGGGCGTTCATGGGTTCGGTGGTGGTGGTGGCGGTGGCGGTTTAGGCAGCCTCCACCA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCTCCCAGATAGAGCGGCGGCGCACTGTAGGGAAGAACCGGGGGACCCGGTCACTACCAAGGACG 240

Qy   241 CCTCCCAGATAGAGCGGCGGCGCACTGTAGGGAAGAACCGGGGGACCCGGTCACTACCAAGGACG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCCCAGATAGAGCGGCGGCGCACTGTAGGGAAGAACCGGGGGACCCGGTCACTACCAAGGACG 300

Qy   301 CAGACCTCTTTTGAGTATCACGCCCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGACCTCTTTTTGAGTATCACGCCCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTGT 360

Qy   361 TGGGATGGTTGGGGTTAGCCATCCATCCATACCGTACCTGCCTGATAGGGGTCCTTGCGAGGGGAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TGGGATGGTTGGGGTTAGCCATCCATCCATACCGTACCTGCCTGATAGGGGTCCTTGCGAGGGGAT 420

Qy   421 CTGGGAGTCTCGTAGACCGTAGCACATGCCCTGTTATTTCTACTCAAACAGTCCTGTACC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTGGGAGTCTCGTAGACCGTAGCACATGCCCTGTTATTTCTACTCAAACAGTCCTGTACC 480

Qy   481 TGCGCCCAGACGCGCAAGACAGAGACGAGGCTTCATATCCTGTGTCATTAAAAC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TGCGCCCAGACGCGCAAGACAGAGACGAGGCTTCATATCCTGTGTCATTAAAAC 540

Qy   481 ATCTGTTGAAAGGGACAACGAGCAAAGCGCAAAGTCCAGCGCGATGCTCGCGCTCGTAA 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 ATCTGTTGAAAGGGACAACGAGCAAAGCGCAAAGTCCAGCGCGATGCTCGCGCTCGTAA 600

Qy   541 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTTCAGGTGCTTTGCCAGC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTTCAGGTGCTTTGCCAGC 660

Qy   601 TCATGGTTGGGACGCCAAGACCCTCGGCATAAGTCTTCGCAATCTTGGAATCTTCTGA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TCATGGTTGGGACGCCAAGACCCTCGGCATAAGTCTTCGCAATCTTGGAATCTTCTGA 720

Qy   661 TTACCCTTGGGGTGGGATTGGTGATGTTACAACTCACACACCTCTAGGAGGCCCGCTGGT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TTACCCTTGGGGTGGGATTGGTGATGTTACAACTCACACACCTCTAGGAGGCCCGCTGGT 780

Qy   721 GGCAGGAGGCGGTCGTCGTTCGACCAGTCTGCCAGATAGTACGCTTGCTGCTGGAGGATGGAGTCAA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GGCAGGAGGCGGTCGTCGTTCGACCAGTCTGCCAGATAGTACGCTTGCTGCTGGAGGATGGAGTCAA 840

Qy   781 CTGGGCTACTGGTGGTTCGGTGTCCACCTTTTTGTGGTATGTCTCGTCATCTTTGGCCTG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CTGGGCTACTGGTGGTTCGGTGTCCACCTTTTTGTGGTATGTCTCGTCATCTTTGGCCTG 900

Qy   841 CTGGGCTACTGGTGGTTCGGTGTCCACCTTTTTGTGGTATGTCTCGTCATCTTTGGCCTG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCCTGTAGTGTGCGGACGAGTGCTGGGGTTCCCGCCAATCCGTACATCCGTACACCCTTCCAA 960

Qy   901 TCCCTGTAGTGTGCGGACGAGTGCTGGGGTTCCCGCCAATCCGTACATCCGTACACCCTTCCAA 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCCTGTAGTGTGCGGACGAGTGCTGGGGTTCCCGCCAATCCGTACATCCGTACACCCTTCCAA 960

Qy   961 CTGCCAGCTACTGGTTGGTTCGGTCATCTTGTCTCCTTCCACTTGCGTGCTCTGCTACAGAGCCTGGTG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CTGCCAGCTAATCAGTTATCATTGTTCTCCTTCCACTTGCGTCTACAGAGCCTGGTG 1020

Qy   1021 TGTGATCTGTGCGGACGAGTGCTGGGGTTCCCGCCAATCCGTACATCTCACACCTTCCAA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TGTGATCTGTGCGGACGAGTGCTGGGGTTCCCGCCAATCCGTACATCTCACACCTTCCAA 1080

Qy   1081 TTGGACTGGCACGGACTCCTTCTTGGCTGGCTGACCACATTGATTTGTTATGGGCGCTCTTGT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TTGGACTGGCACGGACTCCTTCTTGGCTGGCTGACCACATTGATTTGTTATGGGCGCTCTTGT 1140

Qy   1141 GACCTGTGACGCCCTTGACATTGTGAGTTGTCGTGTATTAGTCGTGACTG 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTGTGACGCCCTTGACATTGTGAGTTGTCGTGTATTAGTCGTGACTG 1200

Qy   1201 GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTACCTGA 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTACCTGA 1260

Qy   1261 AGTGCCACTGGAATAGATCCTGGGTTCCTAGGGTTCCTAGGGTTATCGGGTGATGGCCCGGCAAGGT 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGTGCCACTGGAATAGATCCTGGGTTCCTAGGGTTCCTAGGGTTATCGGGTGATGGCCCGGCAAGGT 1320

Qy   1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT 1380

Qy   1381 GTTTAGCAGTGTACACTACCTGGCCGCTTGGCCGCTCTGCTGGGGCAA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GTTTAGCAGTGTACACTACCTGGCCGCTTGGCCGCTCTGCTGGGGCAA 1440

Qy   1441 GTGGTATCAGTTGCTCCTAGCGTTATGTCTTTACATAGAAGCGACCCTCTGGAAACCCCAT 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GTGGTATCAGTTGCTCCTAGCGTTATGTCTTTACATAGAAGCGACCCTCTGGAAACCCCAT 1500

Qy   1501 CAGGGTGCCACTGGATGCTCAATAGCTGAGTTTTCGCCCTTTGATGATACCATGTCC 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTTCGCCCTTTGATGATACCATGTCC 1560

Qy   1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTGCATTTGTTACAGTCCAAAGTGAC 1620

Db 1561 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTTGTTACAGTCCAAAGTGGAC 1620
Db
Qy 1621 CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG 1680
Db 1621 CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG 1680
Qy 1681 TGCGAGGGATGGTATGGTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA 1740
Db 1681 TGCGAGGGATGGTATGGTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA 1740
Qy 1741 TGTGCCATCGTACTGCACTATGGCACTGATGCAGTGTGGAACGACACTCGCACAACACTTA 1800
Db 1741 TGTGCCATCGTACTGCACTATGGCACTGATGCAGTGTGGAACGACACTCGCACAACACTTA 1800
Qy 1801 CGAAGCATGCGGTGTAACACCATGGCTAACAACCGATGCACAACGGCTCAGCCCTGAA 1860
Db 1801 CGAAGCATGCGGTGTAACACCATGGCTAACAACCGATGCACAACGGCTCAGCCCTGAA 1860
Qy 1861 ATTGGCTATATTACAAATACCCTGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC 1920
Db 1861 ATTGGCTATATTACAAATACCCTGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC 1920
Qy 1921 AGGCCATTTGTATTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA 1980
Db 1921 AGGCCATTTGTATTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA 1980
Qy 1981 TTCCACTCTCCTACCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCTGTGGTACG 2040
Db 1981 TTCCACTCTCCTACCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCTGTGGTACG 2040
Qy 2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100
Db 2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100
Qy 2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGTCTTATATTCCGCGCGGTGC 2160
Db 2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGTCTTATATTCCGCGCGGTGC 2160
Qy 2161 TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGTGGGTGTGTGG 2220
Db 2161 TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGTGGGTGTGTGG 2220
Qy 2221 CAGCAAGTATCTTATTTTAGCGTTCTTGTTACTTGTCCCTTGTTTTGTGGGCGCGTTC 2280
Db 2221 CAGCAAGTATCTTATTTTAGCGTTCTTGTTACTTGTCCCTTGTTTTGTGGGCGCGTTC 2280
Qy 2281 TCGTTACCCCTTTGCTCTGTGTTCCCATCCCAGTCGTATCTTGCTATCTCCG 2340
Db 2281 TCGTTACCCCTTTGCTCTGTGTTCCCATCCCAGTCGTATCTTGCTATCTCCG 2340
Qy 2341 TTTGTCTAAAGCTCAAGTAGCTCCTTTCCTTTGCTTTGATTTTCTTCATCTGTTGCTTCCG 2400
Db 2341 TTTGTCTAAAGCTCAAGTAGCTCCTTTCCTTTGCTTTGATTTTCTTCATCTGTTGCTTCCG 2400
Qy 2401 CTGCAGGCTACGTTATGCTGCCCTTTAGGGGTTTGTGCCCAACCAGATTATGACTGGTGCGGACTGCT 2460
Db 2401 CTGCAGGCTACGTTATGCTGCCCTTTAGGGGTTTGTGCCCAACCAGATTATGACTGGTGCGGACTGCT 2460
Qy 2461 AACTTTCTTTTGTTTGCAGGAGCTGCTGCCCAACCAGATTATGACTGGTGCGGACTGCT 2520
Db 2461 AACTTTCTTTTGTTTGCAGGAGCTGCTGCCCAACCAGATTATGACTGGTGCGGACTGCT 2520
Qy 2521 AGTGGCAGGGTTAGTTTTGTGGGCGGCCGGCCGTAACCGTGTCACCGCATAGCTCTGCTTGT 2580
Db 2521 AGTGGCAGGGTTAGTTTTGTGGGCGGCCGGCCGTAACCGTGTCACCGCATAGCTCTGCTTGT 2580
Qy 2581 AGGTCCTTGCCCTCTGGTACGCGTTTAACCCTCCTTGTTTACGCCTTGCTTCAGC 2640
Db 2581 AGGTCCTTGCCCTCTGGTACGCGTTTAACCCTCCTTGTTTACGCCTTGCTTCAGC 2640
Qy 2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTGTAGTAGCATTAGTTGTCAT 2700
Db 2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTGTAGTAGCATTAGTTGTCAT 2700

Db 2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTGTAGTAGCATTAGTTGTCAT 2700
Qy 2701 GTCTCGTTTGGCTTCTTGCTTCTTGCTCTTTGGGCTTGTGCTTTAGTTAGTAACTCCTATCT 2760
Db 2701 GTCTCGTTTGGCTTCTTGCTTCTTGCTCTTTGGGCTTGTGCTTTAGTTAGTAACTCCTATCT 2760
Qy 2761 TTGCAACGTTGGGAGAATTGGTTTGGAACGTTACACTAAGACCGGAGAGGTTTTTCCT 2820
Db 2761 TTGCAACGTTGGGAGAATTGGTTTGGAAAGTTACACTAAGAACCGGAGAGGTTTTTCCT 2820
Qy 2821 TGTGCTCGGTTGTTTCCCCGGTGCGACATATGACGCGGTGGTGACTTTCTGTGTCGTGTCA 2880
Db 2821 TGTGTCGGTTGTTTCCCCGGTGCGACATATGACGCGGTGGTGACTTTCTGTGTGTCA 2880
Qy 2881 CGTAGCTCCTCTCATGTTTAACATCCAGTGCAGCATCGTTCTTGGGACTGACTCTAGGGT 2940
Db 2881 CGTAGCTCCTCTCATGTTTAACATCCAGTGCAGCATCGTTCTTGGGACTGACTCTAGGGT 2940
Qy 2941 TAGGGCCCATAGAAGTGTTCGTGCGTCTCCGGAAAGTGTCATGCTTGGTATTCTCATTATGT 3000
Db 2941 TAGGGCCCATAGAAGTGTTCGTGCGTCTCCGGAAAGTGTCATGCTTGGTATTCTCATTATGT 3000
Qy 3001 TCTTAAGTTTTTCTCTTTAGTGTTTGGTTGAGAATGGTGTGTTTTCTATAAGCACTTGCA 3060
Db 3001 TCTTAAGTTTTTCTCTTTAGTGTTTGGTTGAGAATGGTGTGTTTTCTATAAGCACTTGCA 3060
Qy 3061 TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTCGAAAACTACCCATTTTTCCC 3120
Db 3061 TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTCGAAAACTACCCATTTTTCCC 3120
Qy 3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACGCGTTGGGCGTGTGGGGACAC 3180
Db 3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACGCGTTGGGCGTGTGGGGACAC 3180
Qy 3181 GGTTGATGTTTGCCCGGTTGTTGCCCGTTGTTGCCCGTTGCGTCTCGGCGCGACCTTGTTTTCGCAGGGTTGGCTAT 3240
Db 3181 GGTTGATGTTTGCCCGGTTGTTGCCCGTTGTTGCCCGTTGCGTCTCGGCGCGACCTTGTTTTCGCAGGGTTGGCTAT 3240
Qy 3241 GCCGCCAGATGGGCGTGGGCCATTACCGCACCTTTTACGCGTGCAGTGTCTCTGAACGTGG 3300
Db 3241 GCCGCCAGATGGGCGTGGGCCATTACCGCACCTTTTACGCGTGCAGTGTCTCTGAACGTGG 3300
Qy 3301 CACGCTGTCAGCGATGGCAGTGGTCATGGCGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC 3360
Db 3301 CACGCTGTCAGCGATGGCAGTGGTCATGGCGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC 3360
Qy 3361 TATCTTCAGATTAGGATCTCTCGGCCACTAGCTACATGGGATTTGTTTGTTGTGACAACGTGTT 3420
Db 3361 TATCTTCAGATTAGGATCTCTCGGCCACTAGCTACATGGGATTTGTTTGTTGTGACAACGTGTT 3420
Qy 3421 GTATACTGCTCACCATGGCAGCGAGCAAGAGGGCCGGGTTTGGCTCATCCCACAGGCTCTATACA 3480
Db 3421 GTATACTGCTCACCATGGCAGCGAGCAAGAGGGCCGGGTTTGGCTCATCCCACAGGCTCTATACA 3480
Qy 3481 CCCAATAACCGTTGACGCGGTTGACGCGGTTGACGCGGTCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG 3540
Db 3481 CCCAATAACCGTTGACGCGGTTGACGCGGTTGACGCGGTCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG 3540
Qy 3541 GTCCCTTACTCGGTGCTCTTGCGGGGAGACCAAGGGGGTATCTGGGGTAACACGACTGGGGTC 3600
Db 3541 GTCCCTTACTCGGTGCTCTTGCGGGGAGACCAAGGGGGTATCTGGGGTAACACGACTGGGGTC 3600
Qy 3601 ATTCGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGCGGGGCCCTTCCCAT 3660
Db 3601 ATTCGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGCGGGGCCCTTCCCAT 3660
Qy 3661 GGCTGTTGCCAAGGGTCTTCAGGTGCCCGATTCGTGCTCCTCCGGGCCATGTTATTGG 3720
Db 3661 GGCTGTTGCCAAGGGTCTTCAGGTGCCCGATTCGTGCTCCTCCGGGCCATGTTATTGG 3720
Qy 3721 GATGTTCACCGCTGCTAGAAATTCTGGCCGGTTCAGTCAGTCAGTTAGGCCGTT 3780
Db 3721 GATGTTCACCGCTGCTAGAAATTCTGGCCGGTTCAGTCAGTCAGTTAGGCCGTT 3780

Left column:

```
Qy  3781 GGTGTGTGCTGATGACCATCCCCAGTACACAGCACATGCCACTCCTTGATACAAAACCTAC 3840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781 GGTGTGTGCTGATGACCATCCCCAGTACACAGCACATGCCACTCCTTGATACAAAACCTAC 3840

Qy  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCACTGGCAGCGGCAAGTCAAC 3900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCACTGGCAGCGGCAAGTCAAC 3900

Qy  3901 CAAATTACCACTTTCTTACATGCAGAGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT 3960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 CAAATTACCACTTTCTTACATGCAGAGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT 3960

Qy  3961 GGCTACAACAGCATCAATGCCAAAGTACATGCACGGCACGCGTACCAGATAACTTG 4020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 GGCTACAACAGCATCAATGCCAAAGTACATGCACGGCACGCGTACCAGATAACCAAATTG 4020

Qy  4021 CTATTTTAAATGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGCAT 4080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 CTATTTTAAATGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGCAT 4080

Qy  4081 GTACCTGACCGGAGCAATGTTCCCGGAACTATGATGTAATCATTGTGACGAATGCCATGC 4140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 GTACCTGACCGGAGCAATGTTCCCGGAACTATGATGTAATCATTGTGACGAATGCCATGC 4140

Qy  4141 TACCGGTCAACCACCCGTGTTGGGCATTGGGCTACCACCCGTAATCCCCCTTTGAGGCTACCAAAAAA 4200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 TACCGGTCAACCACCCGTGTTGGGCATTGGGCTACCACCCGTAATCCCCCTTTGAGGCTACCAAAAAA 4200

Qy  4201 TGTTAGGCTAGTGGTTCTTGCACGGCTACCCCCCTGGAGTAATAACGACTGTCTTACTATGGGG 4260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201 TGTTAGGCTAGTGGTTCTTGCACGGCTACCCCCCTGGAGTAATAACGACTGTCTTACTATAGGGG 4260

Qy  4261 CAACATAACTTGAGATTCAATTCAAAAATCCTGAGGGCGCACTATCCCTTTCATGGAAAAAAGAT 4320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261 CAACATAACTTGAGATTCAATTCAAAAATCCCGAGGGCGCACTATCCCTTTCATGGAAAAAAGAT 4320

Qy  4321 TAAGGAGGAAAATCTGAAGAAAGGGGAGCACACTGGATCTGTCTTACTACTAAGGCACACACTG 4380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321 TAAGGAGGAAAATCTGAAGAAAGGGGAGCACACTGGATCTGTCTTACTACTAAGGCACACACTG 4380

Qy  4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTTCTTACTATAGGGG 4440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTTCTTACTATAGGGG 4440

Qy  4441 ATGTGACATCTCAAAAATCCCTGAGCGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441 ATGTGACATCTCAAAAATCCCTGAGCGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500

Qy  4501 TACAGGGTACACTGGGTGACTTTGATTCCGTGTGATGACTGCAGCCTCAGCCTTCAGC 4560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501 TACAGGGTACACTGGGTGACTTTGATTCCGTGTGATGACTGCAGCCTCAGCCTTCAGC 4560

Qy  4561 ATGCCATGGTTGACCTTGACCCTACTTTCACCATGGGGTTCGTGTGTGCGCGGGTTTCAGC 4620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561 ATGCCATGGTTGACCTTGACCCTACTTTCACCATGGGGTTCGTGTGTGCGCGGGTTTCAGC 4620

Qy  4621 AATAGTTAAAGGCCAGCCGTAGGGGCCGCACAGGCGTGGGAGAGCTGGCATATACTACTA 4680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621 AATAGTTAAAGGCCAGCCGTAGGGGCCGCACAGGCGTGGGAGAGCTGGCATATACTACTA 4680

Qy  4681 TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTGTCATCAACAGAGCTCAAACATTGTTGAAGCCTT 4740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681 TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTGTCATCAACAGAGCTCAAACATTGTTGAAGCCTT 4740

Qy  4741 CGACGCAGCCAAGGCCCAAACCTTGGGTATGGTTGTCATCAACAATATTCTGACAC 4800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741 CGACGCAGCCAAGGCCCAAACCTTGGGTATGGTTGTCATCAACAATATTCTGACAC 4800

Qy  4801 CTATCGCACCCCAACCTTGGGTTACCTGCGATAGGAGGAGCAAATTTGGACGAGTGGGCTGATCT 4860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801 CTATCGCACCCCAACCTTGGGTTACCTGCGATAGGAGGAGCAAATTTGGACGAGTGGGCTGATCT 4860
```

Right column:

```
Qy  4861 CTTTTCTCATGGTCAACCCCGAACCTTCATTTGTCAATACTGCAAAAAGAACTGCTGACAA 4920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861 CTTTTCTCATGGTCAACCCCGAACCTTCATTTGTCAATACTGCAAAAAGAACTGCTGACAA 4920

Qy  4921 TTATGTTTGTTGACTGCAGCCAACTACAACTGTGTCATCAGTATGTCTATGCTGCTCC 4980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921 TTATGTTTGTTGACTGCAGCCAACTACAACTGTGTCATCAGTATGGTCTATGCTGCTCC 4980

Qy  4981 CAATGAGCGACCACCACGGTGGCAGGGAGCCGGCGTTGGGAAAAAACCTTGTGGGGTTCTGTG 5040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981 CAATGAGCGACCACCACGGTGGCAGGGAGCCGGCGTTGGGAAAAAACCTTGTGGGGTTCTGTG 5040

Qy  5041 GCGCTTGGACGGCGCTGACGCGCTGTCCTGGCCCAGAGCCCAGCGAGGTGACCAGATACCA 5100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041 GCGCTTGGACGGCGCTGACGCGCTGTCCTGGCCCAGAGCCCAGCGAGGTGACCAGATACCA 5100

Qy  5101 AATGTGCTTCACTGAAGTCAATACTTCTGGGACAGCAGCCCACTCGCTGTGTTGGCGTTGGAGT 5160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101 AATGTGCTTCACTGAAGTCAATACTTCTGGGACAGCAGCCCACTCGCTGTGTTGGCGTTGGAGT 5160

Qy  5161 GGCTATGGCTTATCTAGCCATTGTGACACTTTGTGCGCCGTTGCTGGTC 5220
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161 GGCTATGGCTTATCTAGCCATTGTGACACTTTGTGCGCCGTTGCTGGTC 5220

Qy  5221 TATTACATCAGTCCTACCGGTGCTACCGTGTCGCCCCAGTGGTTGACGAAGAGAAATCGT 5280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221 TATTACATCAGTCCTACCGGTGCTACCGTGTCGCCCCAGTGGTTGACGAAGAGAAATCGT 5280

Qy  5281 GGAGGAGTGTGCATCATTCATCCCCTGGAGGCCATGGTTGCTCAATTGACAAGCTGAA 5340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281 GGAGGAGTGTGCATCATTCATCCCCTGGAGGCCATGGTTGCTCAATTGACAAGCTGAA 5340

Qy  5341 GAGTACAATCACCACAACTAGCTGCAGCTACAACCTTGCTATCATAGAGTATTGCTGTGGTTTAGT 5400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341 GAGTACAATCACCACAACTAGCTGCAGCTACAACCTTGCTATCATAGAGTATTGCTGTGGTTTAGT 5400

Qy  5401 CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCATGTGTTGTCTTTCATTGCTTGCTTTCATTGCTGGG 5460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401 CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCATGTGTTGTCTTTCATTGCTTGCTTTCATTGCTGGG 5460

Qy  5461 CACTTTACCTGACAATCCCTTGTCTTTCATTGCGGGTATTACTAC 5520
         |||||||||||||||||||||||||||||||||||||||||||||
Db  5461 CACTTTACCTGACAATCCCTTGTCTTTCATTGCGGGTATTACTAC 5520

Qy  5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCGTATTATTTGGAGGCGAATTGCGTCCAA 5580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCGTATTATTTGGAGGCGAATTGCGTCCAA 5580

Qy  5581 GCTTACAGACGCTAGGAGGCGGTTGGGTTTGCTTCTTGATGGCGCACTGGGCGCTGCGGGACAGCTCT 5640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581 GCTTACAGACGCTAGGAGGCGGTTGGGTTTGCTTCTTGATGGCGCACTGGGCGCTGCGGGACAGCTCT 5640

Qy  5641 TGGTACATGGACATCGGGGTTTGCTTTGACATCGGGGTGGGGCTAGGCGCTATGCGCCGCCTC 5700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641 TGGTACATGGACATCGGGGTTTGCTTTGACATCGGGGTGGGGCTAGGCGCTATGCGCCGCCTC 5700

Qy  5701 ATCCACTGCTTGCTTGACGCTTAAATGCTTAAATGGCCCACTATGGATCAGCT 5760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701 ATCCACTGCTTGCTTGACGCTTAAATGCTTAAATGGCTGGCCCACTATGGATCAGCT 5760

Qy  5761 TGCTGGTTAGTCTACTTCCGCGTTCAATCCGGCCCAGGAGTTGTGGGCGTCTTGTCAGC 5820
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761 TGCTGGTTAGTCTACTTCCGCGTTCAATCCGGCCCAGGAGTTGTGGGCGTCTTGTCAGC 5820

Qy  5821 TTGTGCAATGTTGCTTTGACAACAGCAGGGCCAGATCACTGGCCAACAACAGACTTCTTAC 5880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821 TTGTGCAATGTTGCTTTGACAACAGCAGGGCCAGATCACTGGCCAACAACAGACTTCTTAC 5880

Qy  5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGACTACTTTATTGCCACTCGTGACATCCG 5940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGACTACTTTATTGCCACTCGTGACATCCG 5940

Qy  5941 CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCCTGGAGTGTCATATCAGCTTGCAT 6000
```

```
XX
AC   AAT00130;
XX
DT   27-AUG-2003 (revised)
DT   02-JUL-1996 (first entry)
XX
DE   Hepatitis GB virus (HGBV) clone GB contig B.
XX
KW   Hepatitis GB virus; HGBV; diagnosis; treatment; vaccine; reagents; non-A;
KW   non-B; non-C; non-D; non-E; clone; GB contig B; tamarin; infected plasma;
KW   lambda phage; cDNA library; ss.
XX
OS   Hepatitis G virus.
XX
FH   Key            Location/Qualifiers
FT   CDS            446..9040
FT                  /*tag= a
XX
PN   WO9521922-A2.
XX
PD   17-AUG-1995.
XX
PF   14-FEB-1995;  95WO-US002118.
XX
PR   14-FEB-1994;  94US-00196030.
PR   13-MAY-1994;  94US-00242654.
PR   29-JUL-1994;  94US-00283314.
PR   23-NOV-1994;  94US-00341185.
PR   23-NOV-1994;  94US-00344190.
PR   27-JAN-1995;  95US-00344557.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Simons JN, Pilot-Matias TJ, Dawson GJ, Schlauder GG, Desai SM;
PI   Leary TP, Muerhoff AS, Erker JC, Buijk SL, Mushahwar IK;
XX
DR   WPI; 1995-293123/38.
DR   P-PSDB; AAR82072.
XX
PT   Non-A, non-B, non-C, non-D, non-E Hepatitis virus reagents - useful for
PT   diagnosis and therapy of hepatitis GB virus.
XX
PS   Example 9; Page 434-447; 661pp; English.
XX
CC   Double stranded hepatitis GB virus (HGBV) DNA obtd. from HGBV infected
CC   tamarin plasma, using standard procedures, was used to prepare a lambda
CC   phage HGBV cDNA library. Clones were rescued from the lambda phage,
CC   searched against a sequence database and found to be unique HGBV
CC   sequences. The clones were then used to assemble the sequences
CC   AAT00129/30 (GB contig A and B) which encode the proteins AAR94345-47
CC   (the 3 possible coding strand reading frames) and AAR82072, respectively.
CC   Reagents which comprise the HGBV DNA, or its protein prods. can be used
CC   for the diagnosis, therapy or in a vaccine to prevent HGBV infection.
CC   (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 9143 BP; 2071 A; 2266 C; 2349 G; 2457 T; 0 U; 0 Other;

Query Match             96.4%;  Score 9059.8;  DB 2;  Length 9143;
Best Local Similarity   99.6%;  Pred. No. 0;
Matches 9096;  Conservative 0;  Mismatches 37;  Indels 4;  Gaps 1;
```

```
Qy  1261  AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT  1320

Qy  1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCACATACGCTATTGCGACTAT  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCACATACGCTATTGCGACTAT  1380

Qy  1381  GTTTAGCAGTGTACACTACCTGGCGCGTTGGCGTCTCGATCTACTACTATGCCTCTCGGGCAA  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  GTTTAGCAGTGTACACTACCTGGCGCGTTGGCGTCTCGATCTACTACTATGCCTCTCGGGCAA  1440

Qy  1441  GTGGTATCAGTTGCTCCTAGCGCTTATGCTTTACATAGAAGCGACCCTCTGGAAACCCCAT  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  GTGGTATCAGTTGCTCCTAGCGCTTATGCTTTACATAGAAGCGACCCTCTGGAAACCCTAT  1500

Qy  1501  CAGGGTGCCCACTGGATGCTCAATAGCTTCAGAAGTCATTTGTTACAGTCCAAGTGGAC  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  CAGGGCTGCCCACTGGATGCTCAATAGCTTCAGAAGTCATTTGCTCGCCTTGATGATACCATGTCC  1560

Qy  1561  TTGCCACTCTTATTGAGTGAGAATGTGTCAGAAGTCATTCCATATCTTGTTACACCCTATATCGCCTATTCGAA  1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  TTGCCACTCTTATTGAGTGAGAATGTGTCAGAAGTCATTCCATATCTTGTTACACCCTATACAATCCCTGG  1620

Qy  1621  CAGGCCTATCACTCTAGAGTATATAACAACTCCATATCTTGGTACCCGGGTTGCTGCGTATTCGAA  1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  CAGGCCTGTCACTCTAGAGTATATAACAATCCATATCTTGGTACCCGGGTTGCTGCCGTATTCGCAA  1680

Qy  1681  TGCCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  TGCCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA  1740

Qy  1741  TGTGCCATCGTACTGCACTATGGGCACTGATGCAGTGTGGAACGACACTCGCAACACTTA  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  TGTGCCATCGTACTGCACTATGGGCACTGATGCAGTGTGGAACGACACTCGCAACACTTA  1800

Qy  1801  CGAAGCATGCGGTGTAACACCATGGCTAACAACCGCACAACGGCTCAGCCCTGAA  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  CGAAGCATGCGGTGTAACACCATGGCTAACAACCGCACAACGGCTCAGCCCTGAA  1860

Qy  1861  ATTGGCTATATTACAAATACCCGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC  1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  ATTGGCTATATTACAATACCCGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC  1920

Qy  1921  AGGCCATTTGTATTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  AGGCCATTTGTATTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA  1980

Qy  1981  TTCCACTCCTACCACCGGAGAGGTGGCTAGGTTGCCCGGTACCCCACCTGTGGTACG  2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  TTCCACTCCTACCACCGGAGAGGTGGCTAGGTTGCCCGGTACCCCACCTGTGGTACG  2040

Qy  2041  TGGTTCTTGGTTACAGTTTCCGCAAGGGTTTCACAGTGATGTGAAAGACCTAGCCACAGG  2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  TGGTTCTTGGTTACAGTTTCCGCAAGGGTTTCACAGTGATGTGAAAGACCTAGCCACAGG  2100

Qy  2101  ATTGATCACCAAAGACAAAGCCTTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC  2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCCGCCACGGGTGC  2160

Qy  2161  TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGGTTGGGGTTGTGTGG  2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGGTTGGGGTTGTGTGG  2220

Qy  2221  CAGCAAGTATCTTATTTTAGCCTACCTCTGTCCCTTTGTTTTGGGCGCGCTTC  2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  CAGCAAGTATCTTATTTTAGCCTACCTCTGTCCCTTTGTTTTGGGCGCGCTTC  2280

Qy  2281  TGGTTACCCTTTGGTTCCAAGCTCGTGTCTCCCATCCCAGTCGTATCTCCAAGTGGATGT  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TGGTTACCCTTTGGTTCCAAGCTCGTGTCTCCCATCCCAGTCGTATCTCCAAGTGGATGT  2340

Qy  2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTGCTTTGATTTTCTTCATCTGTGCTATCTCCG  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy  181   CCTGATGGGCGGTTCATGGGTTCGGTGGTGTGGCGGCCTTTAGGCAGCCTCCACGCCCACCA  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181   CCTGATGGGCGGTTCATGGGTTCGGTGGTGTGGCGGCCTTTAGGCAGCCTCCACGCCCACCA  240

Qy  241   CCTCCCAGATAGAGCGGCGGCACTGTAGGGAGACCGGGGACCGGTCACTACCAAGAGACG  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241   CCTCCCAGATAGAGCGGCGGCACTGTAGGGAGACCGGGGACCGGTCACTACCAAGAGACG  300

Qy  301   CAGACCCTCTTTTGAGTATCACGGCCTCCGGAAGTAGTTGGGCAAGCCCACTTATATGTGT  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301   CAGACCCTCTTTTGAGTATCACGGCCTCCGGAAGTAGTTGGGCAAGCCCACTTATATGTGT  360

Qy  361   TGGGATGGTTGGGGTTAGCCATCCATACGTACTGCCTGATAGGGTCCTTGCGGAGGGGAT  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361   TGGGATGGTTGGGGTTAGCCATCCATACGTACTGCCTGATAGGGTCCTTGCGGAGGGGAT  420

Qy  421   CTGGGAGTCTCGTAGACCGTGACATGCCTGTTATTTCTACTCAAAACAAGTCCTGTACC  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421   CTGGGAGTCTCGTAGACCGTGACATGCCTGTTATTTCTACTCAAAACAAGTCCTGTACC  480

Qy  481   TGCGCCCAGAAGAACAACGCAGAGCAAAGTCAGCGCTTCATATCCTGTGTCTCGGCCTCGTAA  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481   TGCGCCCAGAAGAACAACGCAGAGCAAAGTCAGCGCTTCATATCCTGTGTCTCGGCCTCGTAA  540

Qy  541   ATCTGTTGAAAGGGGACAAAGTCCAGCGCAAAGTCAGCGCTTGGCTCAGGCTGCTTTGCCAGC  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541   ATCTGTTGAAAGGGGACAAAGTCCAGCGCAAAGTCAGCGCTTGGCTCAGGCTGCTTTGCCAGC  600

Qy  601   TTTACAAAATTGCTGTGGTATCCATGATGGCTTGCCAGATAGTACGGTTGCTGGAGGATGGAGTCAA  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601   TTTACAAAATTGCTGTGGTATCCATGATGGCTTGCCAGATAGTACGGTTGCTGGAGGATGGAGTCAA  660

Qy  661   TCATGGTTGGGGGACGCCAAGACCCTCGCCATAGTCTCGCAATCTTGGAATCCTTCTGA  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661   TCATGGTTGGGGGACGCCAAGACCCTCGCCATAGTCTCGCAATCTTGGAATCCTTCTGA  720

Qy  721   TTACCCTTTGGGGTGGATTGGTGATGTTACAACTCACACACACACCTCTAGGCCCGCTGGT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721   TTACCCTTTGGGGTGGATTGGTGATGTTACAACTCACACACACACCTCTAGGCCCGCTGGT  780

Qy  781   GGCAGGAGCGGTCGGTTGCTTGCTGCCAGTGTCTGCCAGATAGTACGGTTGCTGCGGAGGATGGAGTCAA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781   GGCAGGAGCGGTCGGTTGCTTGCTGCCAGTGTCTGCCAGATAGTACGGTTGCTGCGGAGGATGGAGTCAA  840

Qy  841   CTGGGCTACTGGTTGGTTCGGTGTTCGGTGTCCACCCTTTGTGTGGTATGTCTGCTCATCTTTGGGCCTG  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841   CTGGGCTACTGGTTGGTTCGGTGTTCGGTGTCCACCCTTTGTGTGGTATGTCTGCTCATCTTTGGGCCTG  900

Qy  901   TCCCTGTAGTGGGCGGCGGCGGGGTCACTGGGTCCGGAGTGTCACTCCTTGCTGCTGACCAGTACGGTTTTATGGGGCGCTCTGT  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901   TCCCTGTAGTGGGCGGCGGCGGGGTCACTGGGTCCGGAGTGTCACTCCTTGCTGCTGACCAGTACGGTTTTATGGGGCGCTCTGT  960

Qy  961   CTGCCAGCGTAATCGTAATCAGGTTATCTATTGTTCTCCTTCCACTGGCCTACAGGCCTGGTTG  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961   CTGCCAGCGTAATCGTAATCAGGTTATCTATTGTTCTCCTTCCACTGGCCTACAGGCCTGGTTG  1020

Qy  1021  TGTGATCTGCGCGGACGAGTGCGGTGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  TGTGATCTGCGCGGACGAGTGCGGTGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA  1080

Qy  1081  TTGGACTGGCACGGACTCCTTGGCTGACCACATTGATTTGTATGGGGCGTCTTGT  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  TTGGACTGGCACGGACTCCTTGGCTGACCACATTGATTTGTATGGGGCGTCTTGT  1140

Qy  1141  GACCTGTGTGACGCCCTTGACATTGGTGAGTTGGTGTGGTGTATTAGTGGTGCTTCGGTGACTG  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  GACCTGTGTGACGCCCTTGACATTGGTGAGTTGGTGTGGTGTATTAGTGGTGCTTCGGTGACTG  1200

Qy  1201  GCTTGTCAGGCAGCGGCTTATTCACAATAGACCTCAATGAAACTGCTGGTACTTGTTACCTCGGA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GCTTGTCAGGCAGCGGCTTATTCACAATAGACCTCAATGAAACTGCTGGTACTTGTTACCTCGGA  1260

Qy  1261  AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT  1320
```

```
Db  2341 TTTGTCTAAAGCTCAAGTAGCTCCTTTGCTTTGATTTTCTCATCTGTTGCTATCTCCG 2400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2401 CTGCAGGTACGTTATGCTGCCCATGGGTTGTTGTCCCATGGCTTGCGCGGGCTTGCCCCT 2460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 CTGCAGGCTACGTTATGCTGCCCATGGGTTGTTGTCCCATGGCTGCGCGGGCTTGCCCCT 2460

Qy  2461 AACTTTCTTTGTGCAGCAGCTGCTGCCAACCAGATTATGACTGGTGGGTGCGGGTGCGCT 2520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 AACTTTCTTTGTGCAGCAGCTGCTGCCAACCAGATTATGACTGGTGGGTGCGGGTGCGCT 2520

Qy  2521 AGTGGCAGGGTTAGTTTGTGGGGCCGGCCGTAACCGTGTAACCGTGGTCTGTTGCTTGT 2580
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521 AGTGGCAGGGTTAGTTTGTGGGGCCGGCCGTGACCGTGGTCCACGTAGCTCTGCTTGT 2580

Qy  2581 AGGTCCTTGGCCTCTGTAGCCTTTTAACCCTCTGCATTGGTTACGCCTGCTTCAGC 2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581 AGGTCCTTGGCCTCTGTAGCCTTTTAACCCTCTGCATTGGTTACGCCTGCTTCAGC 2640

Qy  2641 TTTTGATACCGGAGATAATTGAGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT 2700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641 TTTTGACACCGGAGATAATTGAGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT 2700

Qy  2701 GTCTCGTTTTGGCTTCTTTGCTCACTTGTTACCTCCGCGTGCTCTAAGACCGGAGAGTTTTCCT 2760
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 GTCTCGTTTTGGCTTCTTTGCTCACTTGTTACCTCCGCGTGCTCTAAGACCGGAGAGTTTTCCT 2760

Qy  2761 TTGGCAACGTTGGGAGAATTGTTTGTGGAACGTTACACTAAGACCGGAGAGTTTCTCCT 2820
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 TTGGCAACGTTGGGAGAATTGTTTGTGGAACGTTACACTAAGACCGGAGAGTTTCTCCT 2820

Qy  2821 TGTGCTGGTTGTTTCCCCGTGCGACATATGACGGCGTGGTGACTTCTGTGTGTGTCA 2880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821 TGTGCTGGTTGTTTCCCCGTGCGACATATGACGGCGTGGTGACTTCTGTGTGTGTCA 2880

Qy  2881 CGTAGCTCTTCTCATGTTAACATCCAGTGCAGCATCGTCTTCTTGGGACTGACTCTAGGGT 2940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881 CGTAGCTCTTCTCATGTTAACATCCAGTGCAGCATCGTCTTCTTGGGACTGACTCTAGGGT 2940

Qy  2941 TAGGGCCCATAGAATGTTGGTGCTCTCGGAAAGTGTCATGTTGGTATTCTCATTATGT 3000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 TAGGGCCCATAGAATGTTGGTGCTCTCGGAAAGTGTCATGTTGGTATTCTCATTATGT 3000

Qy  3001 TCTTAAGTTTTCCCTCTAGTGTTGGTGAATGTCTGAAATGTGTGTTTTCTATAAGCACTTGCA 3060
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 TCTTAAGTTTTCCCTCTAGTGTTGGTGAATGTCTGAAATGTGTGTTTTCTATAAGCACTTGCA 3060

Qy  3061 TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCGGCGTGTGGGCACAC 3120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCGGCGTGTGGGGACAC 3120

Qy  3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGAAGACGCTTGGCGTGTGGGGACAC 3180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGAAGACGCTTGGCGTGTGGGGACAC 3180

Qy  3181 GGTTGATGGTTTGCCCGTTGTTGCGCGGGTCTCGGCGGACCTTGTTTTCGCAGGGTTGGCTAT 3240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 GGTTGATGGTTTGCCCGTTGTTGCGCGGGTCTCGGCGGACCTTGTTTTCGCAGGGTTAGCTAT 3240

Qy  3241 GCCGCCAGATGGGTGGGCCATTACGGGCGACACCTTTACGCTGCAGTGTCTCTGAACGTGG 3300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 GCCGCCAGATGGGTGGGCCATTACGGGCGACACCTTTACGCTGCAGTGTCTCTGAACGTGG 3300

Qy  3301 CACGCTGTCAGCGATGGCAGCTGGTCATGTATAGACCCCGAACTTGGACTCGAAC 3360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 CACGCTGTCAGCGATGGCAGCTGGTCATGTATAGACCCCGAACTTGGACTCGGAAC 3360

Qy  3361 TATCTTCAGATTAGGATCTCTGCCACTAGCTACATGGGATTGTTGTGTGACAAGCGTGTT 3420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 TATCTTCAGATTAGGATCTCTGCCACTAGCTACATGGGATTGTTGTGTGACAAGCGTGTT 3420

Qy  3421 GTATACTGCTCACCATGGCAGCAAGGGGGCGCCAAGGGGCCGGTTGGCTCATCCCACAGGCTCTATACA 3480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421 GTATACTGCTCACCATGGCAGCAAGGGGGCGCCAAGGGGCCGGTTGGCTCATCCCACAGGCTCCATACA 3480
```

```
Qy  3481 CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACACCATGTGGAGCTGG 3540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481 CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACACCATGTGGAGCTGG 3540

Qy  3541 GTCCCTACTCGGTCTGTCTTGCGGGGAGACCAAGGGTATCTGTAACGACTGGGGTC 3600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541 GTCCCTACTCGGTCTGTCTTGCGGGGAGACCAAGGGTATCTGTAACGACTGGGGTC 3600

Qy  3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTATTGTGTGTGTGCGGGGCCCTTCCCAT 3660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTATTGGTGTGTGTGCGGGGCCCTTCCCAT 3660

Qy  3661 GGCTGTTGCCAAGGGTTCTTCAGGTGCCCGATTCTGTGCTCGTCCTCCCGGCATGTTATTGG 3720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661 GGCTGTTGCCAAGGGTTCTTCAGGTGCCCGATTCTGTGCTCGTCCTCCCGGCATGTTATTGG 3720

Qy  3721 GATGTTCACCGGTCTGCTAGAAATTCTCAGTGGGTTCAGTCAGTCATTGATACAAAACCTAC 3780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721 GATGTTCACCGGTCTGCTAGAAATTCTCAGTGGGTTCAGTCAGTCATTGATACAAAACCTAC 3780

Qy  3781 GGTGTGTGCTGATACCATCCCAGTACACAGCACATGCCACTCTTGATAATGCGCAGCGGCAAGTCAAC 3840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781 GGTGTGTGCTGATACCATCCCAGTACACAGCACATGCCACTCTTGATAATGCGCAGCGGCAAGTCAAC 3840

Qy  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCACTGGCAGCGGCAAGTGT 3900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCACTGGCAGCGGCAAGTCAAC 3900

Qy  3901 CAAATTACCACTTCTTACATGCAGGAGAAGTATGAGTCTTGGTCCTAAATCCCAGTGT 3960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 CAAATTACCACTTCTTACATGCAGGAGAAGTATGAGTCTTGGTCCTAAATCCCAGTGT 3960

Qy  3961 GGCTACAACAGCCATCAATGCCAAAGTACATGCCAAAGTACACGCGACGGCGAATCCCAAATTG 4020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 GGCTACAACAGCCATCAATGCCAAAGTACATGCCAAAGTACACGCGACGGCGTGAATCCAAATTG 4020

Qy  4021 CTATTTTAATGGCAAATGTACCAACACACAGGGGCTTCACTTACGTACGACAGCACATATGGCAT 4080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 CTATTTTAATGGCAAATGTACCAACACACAGGGGCTTCACTTACGTACGACAGCACATATGGCAT 4080

Qy  4081 GTACCTGACCGGAGCATGTTCCCGGAACTATGATGTAATCATTGTGACGAATGCCATGC 4140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 GTACCTGACCGGAGCATGTTCCCGGAACTATGATGTCATCATTGTGACGAATGCCATGC 4140

Qy  4141 TACCGATGCAACCACCGTGTTGGGCATTGGAAAGGTCCATCCAACCACCACATGC 4200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 TACCGATGCAACCACCGTGTTGGGCATTGGAAAGGTTCAACGAAGCTCCATCCAACCACCACATGC 4200

Qy  4201 TGTTAGGCTAGTGTCTTGCCACCGGGCTAATCCTACCCCCTGGAGTAATCCTACCACCACATGC 4260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201 TGTTAGGCTAGTGTCTTGCCACCGGGCTAATCCTACCCCCTGGAGTAATCCTACCACCACATGC 4260

Qy  4261 CAACATAACTGAGATCAATTAACCGATGAGGCACTATCCCTTTCATGGAAAAAAGAT 4320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261 CAACATAACTGAGATCAATTAACCGATGAGGCACTATCCCTTTCATGGAAAAAAGAT 4320

Qy  4321 TAAGGAGGAAAATCTGAAGAAAGGAGACACCTTATCTTTGAGGCTACCCAAAAAACACTG 4380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321 TAAGGAGGAAAATCTGAAGAAAGGAGACACCTTATCTTTGAGGCTACCCAAAAAACACTG 4380

Qy  4381 TGATGAGCTTGCTAAACAGCAGTAGCTCGAAAGGGAATAACAGCTGTCTTCTTACTAGGGG 4440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381 TGATGAGCTTGCTAAACAGCAGTAGCTCGAAAGGGAATAACAGCTGTCTTCTTACTAGGGG 4440

Qy  4441 ATGTGACATCTCAAAAATCCCTGAGGGCGACTGTGTGTAGTAGTGCCACTGATGCCTTGTG 4500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441 ATGTGACATCTCAAAAATCCCTGAGGGCGACTGTGTGTAGTAGTGCCACTGATGCCTTGTG 4500

Qy  4501 TACAGGGTACACTGGTACTCCGTGTGATGTGATGACTGCAGCCTCATGGTAGAAGGCAC 4560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501 TACAGGGTACACTGGTACTCCGTGTGATGTGATGACTGCAGCCTCATGGTAGAAGGCAC 4560
```

```
Qy  4561  ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTCGTCGTGTGTGCGGGGTTTCAGC  4620
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561  ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTCGTCGTGTGTGCGGGGTCTCAGC  4620

Qy  4621  AATAGTTAAAGGCCAGCGTAGGGCCCGTAGGGGCCGCACCCCGTGGGAGAGCTGCATATACTACTA  4680
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621  AATAGTTAAAGGCCAGCGTAGGGCCCGTAGGGGCCGCACCCCGTGGGAGAGCTGCATATACTACTA  4680

Qy  4681  TGTAGACGGGAGTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGGCTT  4740
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681  TGTAGACGGGAGTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGGCTT  4740

Qy  4741  CGACGCAGCCAAGGCATGGTATGGTTTGTCATCAACAGAGCTCAAACTATTCTGGACAC  4800
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741  CGACGCAGCCAAGGCATGGTATGGTTTGTCATCAACAGAGCTCAAACTATTCTGGACAC  4800

Qy  4801  CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT  4860
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801  CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT  4860

Qy  4861  CTTTTCTATGGTCAACCCGAACCTTCATTGTCAATACTGCAAAAGAAACTGCTGACAA  4920
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861  CTTTTCTATGGTCAACCCGAACCTTCATTGTCAATACTGCAAAAGAAACTGCTGACAA  4920

Qy  4921  TTATGTTTGTTGACTGCAGCCAACTACAACTGTCATCAGTATGGCTATGTCGCTCC  4980
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921  TTATGTTTGTTGACTGCAGCCAACTACAACTGTCATCAGTATGGCTATGTCGCTCC  4980

Qy  4981  CAATGACGCACCACGGTTGGCAGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG  5040
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981  CAATGACGCACCACGGTTGGCAGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG  5040

Qy  5041  GCGCTTGGACGGCGCTGACGCTCCTGGCCCAGCGCCCAGCGCCAGAGCGTGACCAGATACCA  5100
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041  GCGCTTGGACGGCGCTGACGCTCCTGGCCCAGCGCCCAGCGCCAGAGCGTGACCAGATACCA  5100

Qy  5101  AATGTGCTTCACTGAAGTCAATACTTCGGAGCAGCCCACTCGCTGTTGGCCGTTGGGAGT  5160
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101  AATGTGCTTCACTGAAGTCAATACTTCGGAGCAGCCCACTCGCTGTTGGCCGTTGGGAGT  5160

Qy  5161  GGCTATGGCTTATCTCAGCCATTGACACTTTGGCGCCACTTTGCGCGTCGTGTGCTGCTGTC  5220
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161  GGCTATGGCTTATCTCAGCCATTGACACTTTGGCGCCACTTTGCGCGTCGTGTGCTGCTGTC  5220

Qy  5221  TATTACATCAGTCCCTACCGGTGCTACTGTCGCCGCCCCAGTGGTTGACGAAGAAGAAATCGT  5280
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221  TATTACATCAGTCCCTACCGGTGCTACTGTCGCCGCCCCAGTGGTTGACGAAGAAGAAATCGT  5280

Qy  5281  GGAGGAGTGTGCATCATTCCCTTGAGGCCATGGTTGCTCGCAATTGACAAGCTGAA  5340
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281  GGAGGAGTGTGCATCATTCCCTTGAGGCCATGGTTGCTCGCAATTGACAAGCTGAA  5340

Qy  5341  GAGTACAATCACCACAACTAGTCTTTCTTTCACATTGGAAAACCGCCCTTGAAAAACTTAACAC  5400
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341  GAGTACAATAACCACAACTAGTCTTTCTTTCACATTGGAAAACCGCCCTTGAAAAACTTAACAC  5400

Qy  5401  CTTTCTTGGGCCTCATGGGCCTCACAAATCCTTGCTCATCATAGAGTATTGCTGTTGGTTTAGT  5460
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401  CTTTCTTGGGCCTCATGGGCCTCACAAATCCTTGCTCATCATAGAGTATTGCTGTTGGTTTAGT  5460

Qy  5461  CACTTTACCTGACAATCCCTTTGCATCATGCGTGTTTGCTTTCATTGGGGGTATTACTAC  5520
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461  CACTTTACCTGACAATCCCTTTGCATCATGCGTGTTTGCTTTCATTGGGGGTATTACTAC  5520

Qy  5521  CCCACTACCTCACAAGATCAAGATCAAAATGTTCCGTGTCATTATTGGAGGCGCAATTGCGTCAA  5580
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521  CCCACTACCTCACAAGATCAAGATCAAAATGTTCCGTGTCATTATTGGAGGCGCAATTGCGTCAA  5580

Qy  5581  GCTTACAGACGTGAGGGCCACTGGGCGTTCAGGCCGTTCAATGGCGGGGGCTAGAGGCGGACAGCTCT  5640
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581  GCTTACAGACGCTAGAGGCCACTGGGCGTTCAGGCCGTTCAATGGCGGGGGCTAGAGGCGGACAGCTCT  5640

Qy  5641  TGGTACATGGACATCGGTGGGTTTTGTCTTTGACATGGGTCGTCGTGCCGCCTC  5700
                ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  5641  TGGTACATGGACATCGGTGGGTTTTGTCTTTGACATGGGTCGTCGTGCCGCCTC  5700

Qy  5701  ATCCACTGCTTGCTGCTTGTTAAATGCTTGATGGGTGAGTGGCGCCACTATGGATCAGCT  5760
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701  ATCCACTGCTTGCTGCTTGTTAAATGCTTGATGGGTGAGTGGCGCCACTATGGATCAGCT  5760

Qy  5761  TGCTGGTTTAGTCTACTCCGGTCAATCCGGCCGCAGGAGTGTGGGCGTCTTGTCAGC  5820
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761  TGCTGGTTTAGTCTACTCCGGTCAATCCGGCCGCAGGAGTGTGGGCGTCTTGTCAGC  5820

Qy  5821  TTGTGCAATGTTTGCTTTGACAACAGCAGGCCCAGATCACTCGGCCCAACAGACTTCTTAC  5880
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821  TTGTGCAATGTTTGCTTTGACAACAGCAGGCCCAGATCACTCGGCCCAACAGACTTCTTAC  5880

Qy  5881  TATGCTTGCTAGGAGCAACACTGTATGTAGGCATCTAATGAGTACTTTATTGCCACTCCGTGACATCCG  5940
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881  TATGCTTGCTAGGAGCAACACTGTATGTAGGCATCTAATGAGTACTTTATTGCCACTCCGTGACATCCG  5940

Qy  5941  CAGGAAGATACTGGGCATTCTGAGGCATCTACCCCCCTGGAGTGTCATATCAGCTTGCAT  6000
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941  CAGGAAGATACTGGGCATTCTGAGGCATCTACCCCCCTGGAGTGTCATATCAGCTTGCAT  6000

Qy  6001  CCGTTGGCTCCACACCCGACGGAGGATGATTGGGCGCCTCATTGCTTGTTGCTTTGGGTCTCC  6060
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001  CCGTTGGCTCCACACCCGACGGAGGATGATTGGGCGCCTCATTGCTTGTTGCTTTGGGTCTCC  6060

Qy  6061  TTGGCAGTATGTGTCAATTCTTGTGATTGCTTTAATGTCCTTAAAGCTGGAGTTCA  6120
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061  TTGGCAGTATGTGTCAATTCTTGTGATTGCTTTAATGTCCTTAAAGCTGGAGTTCA  6120

Qy  6121  GAGCATGGTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGTGCTGCCAGAAGGGGTACAAGGGCCC  6180
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121  GAGCATGGTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGTGCTGCCAGAAGGGGTACAAGGGCCC  6180

Qy  6181  CTGGATTGGATCAGGTAGTGATGCTCAAGCACGCGTCCATGCGGTGTCGAACTCATCTTTTC  6240
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181  CTGGATTGGATCAGGTAGTGATGCTCAAGCACGCGTCCATGCGGTGTCGAACTCATCTTTTC  6240

Qy  6241  TGTTGAGAATGGTTTGCAAAGCTTTACAAAGGACCCAGAACTTGTTCAAATTACTACTGGAG  6300
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241  TGTTGAGAATGGTTTGCAAAGCTTTACAAAGGACCCAGAACTTGTTCAAATTACTACTGGAG  6300

Qy  6301  AGGGGCTGTTCCAGTCAACGCTGTGTGGTGGTCGGCTGGTCAGCTCGACCGGACCCAACTGATTG  6360
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301  AGGGGCTGTTCCAGTCAACGCTGTGTGGTGGTCGGCTGGTCAGCTCGACCGGACCCAACTGATTG  6360

Qy  6361  GACTAGTCCTTGCTCAATTATGGCGTTAGGGACTACTGTAAATATGGGAGAAAATGGGAGA  6420
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361  GACTAGTCCTTGCTCAATTATGGCGTTAGGGACTACTGTAAATATGGGAGAAAATGGGAGA  6420

Qy  6421  TCACATTTTTGTTACAGCAGTATCCTCCAAATGTCTGTTTCACCCAGGTGCCCCCAAC  6480
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421  TCACATTTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCCAAC  6480

Qy  6481  CTTGAGAGCTGCGAGTGGCCGTGACGGGCGTGGCAGGTTCAGTGTTATCTAGGTGAGCCCAA  6540
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481  CTTGAGAGCTGCGAGTGGCCGTGACGGGCGTGGCAGGTTCAGTGTTATCTAGGTGAGCCCAA  6540

Qy  6541  AACTCCTTGCCTTGGACGACCATCTGCTTGCTGTCTTGCTCGTTACGGTAAGGGTAAAACTGTTAA  6600
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541  AACTCCTTGCCTTGGACGACCATCTGCTTGCTGTCTTGCTCGTTACGGTAAGGGTAAAACTGTTAA  6600

Qy  6601  GCTTCCCTTCCGCGTTGACGGGTGCGCCATGCCATGGTGCGCCATGCCACACCTCAATTTCGTGA  6660
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601  GCTTCCCTTCCGCGTTGACGGGTGCGCCATGCCACACCTCAATTTAATTTCGTGA  6660

Qy  6661  TGCACTTGAGACAAATGACTGTAATTCCACAAACAACACTCCTAGTGATGGCCGCAGT  6720
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661  TGCACTTGAGACAAATGACTGTAATTCCATAAACAACACTCCTAGTGATGGCCGCAGT  6720

Qy  6721  GTCCGCTCTTGTTCAAACAGGAGTTGCCGGCGTACATGGACATCGGTGGGTTTTC  6780
                ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  6721 GTCCGCTCTTGTTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC 6780

Qy  6781 AGCTGGGCTTGACACACCACCAAACTGCCAGCCCCTCCATCGAAGAGTAGTGGTAAGAAA 6840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781 AGCTGGGCTTGACACACCACCAAACTGCCAGCCCCTCCATCGAAGAGTAGTGGTAAGAAA 6840

Qy  6841 GCGGCAGTTCCGGGCAAGAACTGGTTCGCTTGCTTACCTTGCTTGCCTCCCCCTCCGAGATCCGTCCC 6900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841 GCGGCAGTTCCGGGCAAGAACTGGTTCGCTTGCTTACCTTGCTTGCCTCCCCCTCCGAGATCCGTCCC 6900

Qy  6901 AGGAGTGTCATGTCCTGAAAGCCTGCAACGAAGCCGTTAGAACGTCCTTCAAACCT 6960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901 AGGAGTGTCATGTCCTGAAAGCCTGCAACGAAGCCGTTAGAACGTCCTTCAAACCT 6960

Qy  6961 CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGCCATGCCGATGCCCGTTGGGAGCGGGTGA 7020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961 CCCTTCTTCACCACCTGTTCTACAGTTGGCCATGCCATGCCGATGCCCGTTGGGAGCAGTGA 7020

Qy  7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAACAGGCGGAGGCCCTGATGA 7080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAACAGGCGGAGGCCCTGATGA 7080

Qy  7081 TTTACCCAGTTACCCTCCCAAAAAAGAGGTCTCTGAATGGTCAGACGAAAGTTGGTCGAC 7140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081 TTTACCCAGTTACCCTCCCAAAAAAGAGGTCTCTGAATGGTCAGACGAAAGTTGGTCAAC 7140

Qy  7141 GGCTACAACCGCTTCCAGCTTACTGGGCCCCGTACCCTAAGATACGGGGAAGGA 7200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141 GACTACAACCGCTTCCAGCTTACTGGGCCCCGTACCCTAAGATACGGGGAAAGGA 7200

Qy  7201 TTCCACTCAGTCAGCCCCGCCAAACGGCTACACCTGGAATGGTGCACCAGAAGAGTTG 7260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201 TTCCACTCAGTCAGCCCCGCCAAACGGCTACACCTGGAATGGTGCACCAGAAGAGTT 7260

Qy  7261 TTCGTGCAGCATGAGCTGAGACTGGGATGCGCGGGATCGCGGCCATCACTAGTTG 7320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261 TTCGTGCAGCATGAGCTGAGACTGGGATGCGCGGGATCGCGGCCATCACTAGTTG 7320

Qy  7321 TCTGTCTGCAACTCGGGCCATCATTAGTGGTTCCTCAAACAAAAAGTCACTATTAATAGACAACCTCT 7380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7321 TCTGTCTGCAACTCGGGCCATCATTAGTGGTTCCTCAAACAAAAAGTCACTATTAATAGACAACCTCT 7380

Qy  7381 GACTGAGCCGCGGGATGCGGGAGCTTAGAAACAAGCAAGCTTAATAGAAAACAAAAAGTCACTAT 7440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381 GACTGAGCCGCGGGATGCGGGAGCTTAGAAACAAGCAAGCTTAATAGAAAACAAAAAGTCACTAT 7440

Qy  7441 GTTCCCCATCATACCACCAGTTGGCTCGAGCCAGTCATTATCGGCAAATGTGAT 7500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441 GTTCCCCATCATACCACCAGTTGGCTCGAGCCAGTCATTATCGGCAAATGTGAT 7500

Qy  7501 CGGTGTCATGTGGACTATGATGAAGTAGCAGCTCACACGCCCTCTGCTTCGTAAGTC 7560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7501 CGGTGTCATGTGGACTATGATGAAGTAGCAGCTCACACGCCCTCTGCTTCGTAAGTC 7560

Qy  7561 CCACATCACTGCCTTCGGGGCACTCGATGTTGCTTAGGCTGTTCT 7620
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  7561 CCACATCACTGCCTTCGGGGCACTCGATGTTGCTTAGGCTGTTCT 7620

Qy  7621 GGACTTGCAGAAGTGTCGAGGCAGTGAGATAACCGAGTCATTATCGGCAAATCTGAT 7680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7621 GGACTTGCAGAAGTGTCGAGGCAGTGAGATAACCGAGTCATTATCGGCAAATCTGAT 7680

Qy  7681 AGTTCCAAAGGAGGAGGTCTTCGTCGAGAGAAGAGAGCAAAGAAAGAAACCCCCAAG 7740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7681 AGTTCCAAAGGAGGAGGTCTTCGTCGAGAGAAGAGAGCAAAGAAAGAAACCCCCAAG 7740

Qy  7741 GCTTATCTCGTACCCCCACCTTGAAATGAGATGTGTTGAAGAAGATGTCACTACGGGTCAGGT 7800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7741 GCTTATCTCGTACCCCCACCTTGAAATGAGATGTGTTGAAGAAGATGTCACTACGGGTCAGGT 7800

Qy  7801 TGCTCCTGACGTAGTTAAAGCTCATGGGAATGCGTACGGGTTTGTCGACCCCACGTAC 7860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7801 TGCTCCTGACGTAGTTAAAGCTCATGGGAATGCGTACGGGTTTGTCGACCCCACGTAC 7860
```

```
Db  7861 CCGTGTCAAGGTCTCTGTTGTCGATGTGGTCACCCGATGCAGTCGGAGCCACATGCGATAC 7920

Qy  7861 CCGTGTCAAGGTCTCTGTTGTCGATGTGGTCACCCGATGCAGTCGGAGCCACATGCGATAC 7920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7861 CCGTGTCAAGGTCTCTGTTGTCGATGTGGTCACCCGATGCAGTCGGAGCCACATGCGATAC 7920

Qy  7921 AGTGTGTTTGACAGTACCATCACACCCGGAGATATCATGGTGGAGACAGACATCTACTC 7980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7921 AGTGTGTTTGACAGTACCATCACACCCGGAGATATCATGGTGGAGACAGACATCTACTC 7980

Qy  7981 AGCAGCTAAACTCAGTGACCAACACCACCAGACTCACACCATTGGCAGGAGCAGTTATA 8040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7981 AGCAGCTAAACTCAGTGACCAACACCACCAGACTCACACCATTGGCAGGAGCAGTTATA 8040

Qy  8041 CGCTGGACGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGGTAGGTCTAGGTC 8100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8041 CGCTGGACGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGGTAGGTGTAGGTC 8100

Qy  8101 TTCCGGCGTCTATATCTCAAGTCAAGTGAAGAAGAACCCTCGCTTCCTTATTTGCGGCGCTGCTGTCTTTGCTAG 8160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8101 TTCCGGCGTCTATATCTCAAGTCAAGTGAAGAAGAACCCTCGCTTCCTTATTTGCGGCGCTGCTGTCTTTGCTAG 8160

Qy  8161 TGCAGCCGAACAGGCTGGCATGAAGACAATCTCGCTTCCTATTTGCGGCGATGATTGCAC 8220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161 TGCAGCCGAACAGGCTGGCATGAAGACAATCTCGCTTCCTATTTGCGGCGATGATTGCAC 8220

Qy  8221 CGTAATTTGGAAGAGCCCGGAGCAGATGCAGAACAAGCAATGCGTGTCTTTGCTAG 8280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8221 CGTAATTTGGAAGAGCCCGGAGCAGATGCAGAACAAGCAATGCGTGTCTTTGCTAG 8280

Qy  8281 CTGGATGAAGGTGATGGGTGCACCAAGAGATTGTGCTCAACCCAAATACAGTTTGGA 8340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281 CTGGATGAAGGTGATGGGTGCACCAAGAGATTGTGCTCAACCCAAATACAGTTTGGA 8340

Qy  8341 AGAATTAACATCATGTCTCATCAAATGTTACCTCTGGATTAACAAAAGTGGCAAGCCTTA 8400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8341 AGAATTAACATCATGTCTCATCAAATGTTACCTCTGGATTAACAAAAGTGGCAAGCCTTA 8400

Qy  8401 CTACTTTCTTACAAGAGATCCTGCGTATCCCCCTTGGGCAGTGCTCTCGCCGAGGGCTCTGTAG 8460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8401 CTACTTTCTTACAAGAGATCCTGCGTATCCCCCTTGGGCAGTGCTCTCGCCGAGGGCTCTGTAG 8460

Qy  8461 ATACAACCCAGTGCGTCGCGTGGATTGGGTATCTAATACATCAATACACTACTCCCGA 8520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8461 ATACAACCCAGTGCGTCGCGTGGATTGGGTATCTAATACATCAATACACTACTCCCGA 8520

Qy  8521 TAGCCGTGTGTTGGCTGTGTCCATTTCATGGGAAAAATATATACGGTGCCTGTAGAAGATCGCCCAG 8580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8521 TAGCCGTGTGTTGGCTGTGTCCATTTCATGGGAAAAATATATACGGTGCCTGTAGAAGATCGCCCAG 8580

Qy  8581 GACTGTGACCTTTGACTGTGATGGGAAAAATTATACGGTGCCTGTAGAAGATCGCCCAG 8640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8581 GACTGTGACCTTTGACTGTGATGGGAAAAATTATACGGTGCCTGTAGAAGATCGCCCAG 8640

Qy  8641 CATCATTGCTGTGGTGTGACGTATTGGAGGCTTTCTCGGTGTGTCGCTACCACCAACGCTGA 8700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8641 CATCATTGCTGTGGTGTGACGTATTGGAGGCTTTCTCGGTGTGTCGCTACCACCAACGCTGA 8700

Qy  8701 GATCCTCAGAGTTCCCAATCACTAACACATGACACACCTGCGAGCCTGGCG 8760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8701 GATCCTCAGAGTTCCCAATCACTAACACATGACACACCTGCGAGCCTGGCG 8760

Qy  8761 AAAGAAAGCCAGGGCGGCCAAGAGGCGCCAAGAGAGGCGGCGAGCACACCAGAAATT 8820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8761 AAAGAAAGCCAGGGCGGCCAAGAGGCGCCAAGAGAGGCGGCGAGCACACCAGAAATT 8820

Qy  8821 GGCTCGCTTCCTTCTGGCATGTCTACATCTACAGACCTCTACAGATTGGATAAGACGAG 8880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8821 GGCTCGCTTCCTTCTGGCATGTCTACATCTACAGACCTCTACAGATTGGATAAGACGAG 8880

Qy  8881 CGTGGCTCGGTACCACCACTTTCAATTATTGTGATGTTTACTCCCCGGAGGGGATGTGTT 8940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8881 CGTGGCTCGGTACCACCACTTTCAATTATTGTGATGTTTACTCCCCGGAGGGGATGTGTT 8940
```

```
Qy   8941 TATTACCACCAGAGAAGAGATTGCAGAAGTTCCTTGTGAAGTATTGGCTGTCATTGTTTT 9000
          ||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db   8941 TGTTACCACCAGAGAAGAGATTGCAGAAGTTCTTGTGAAGTATTTGGCTGTCATTGTTTT 9000

Qy   9001 TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATCAGTGAACCCCCAAATTCAAAATTAA  9060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9001 TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATCAGTGAACCCCCAAATTCAAAATTAA  9060

Qy   9061 CTAACAG----TTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCCAACCAGGGGAGACCCC 9116
          ||||||    |||||||||||||||||||||||||||||||||||||||||||||||
Db   9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCCAACCAGGGGAGACCCC 9120

Qy   9117 GGGCTTAACGACCCCGC 9133
          |||||||||||||||
Db   9121 GGGCTTAACGACCCCGC 9137

RESULT 7
AAA55376
ID   AAA55376 standard; DNA; 9143 BP.
XX
AC   AAA55376;
XX
DT   06-AUG-2003 (revised)
DT   30-AUG-2000 (first entry)
XX
DE   Hepatitis GB virus nucleotide sequence SEQ ID NO:390.
XX
KW   Hepatitis GB virus; HGBV; diagnosis; therapeutic; immunogenic; infection;
KW   detection; characterisation; hepatitis; ds.
XX
OS   Hepatitis GB virus.
XX
PN   US6051374-A.
XX
PD   18-APR-2000.
XX
PF   07-JUN-1995;   95US-00488445.
XX
PR   14-FEB-1994;   94US-00196030.
PR   13-MAY-1994;   94US-00242654.
PR   29-JUL-1994;   94US-00283314.
PR   23-NOV-1994;   94US-00344185.
PR   23-NOV-1994;   94US-00344190.
PR   30-JAN-1995;   95US-00377557.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Dawson GJ, Leary TP, Muerhoff AS, Pilot-Matias TJ, Buijk SL;
PI   Mushahwar IK, Simons JN, Desai SM, Erker JC, Schlauder GG;
XX
DR   WPI; 2000-338307/29.
XX
PT   Detecting target hepatitis GB virus nucleic acid in a test sample
PT   suspected of containing HGBV comprises reacting the test sample the HGBV
PT   polynucleotide probe and detecting the complex that contains target HGBV.
XX
PS   Example 9; Col 409-418; 369pp; English.
XX
CC   The present invention describe a method for detecting target hepatitis GB
CC   virus (HGBV) nucleic acid (THN) in a test sample (T) suspected of
CC   containing HGBV. The method involves reacting (T) with a HGBV
CC   polynucleotide probe (I) containing 15 contiguous nucleotides, and which
CC   selectively hybridises to the HGBV genome or its full complement, and
CC   detecting the complex that contains THN, indicating the presence of
CC   target HGBV. The method is used for detecting target HGBV nucleic acid in
CC   the test sample suspected of containing HGBV and for characterisation of
CC   newly ascertained etiological agent of non-A, non-B, non-C, non-D and non
CC   -E hepatitis causing agents collectively termed as hepatitis GB virus.
CC   AAA55270 to AAA55489 and AAB08985 to AAB09480 represent nucleotide and
CC   protein sequences used in the exemplification of the present invention.
CC   (Updated on 06-AUG-2003 to correct OS field.)
XX
```

```
SQ   Sequence 9143 BP; 2071 A; 2266 C; 2349 G; 2457 T; 0 U; 0 Other;

     Query Match               96.4%;    Score 9059.8;  DB 3;  Length 9143;
     Best Local Similarity     99.6%;    Pred. No. 0;
     Matches 9096;  Conservative  0;  Mismatches 37;  Indels 4;  Gaps  1;

Qy     1 ACCACAAACACTCCAGTTTGTTACACTCCGTAGAATGCTCCTGGAGCACCCCCTAG   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ACCACAAACACTCCAGTTTGTTACACTCCGTAGAATGCTCCTGGAGCACCCCCTAG   60

Qy    61 CAGGGCGTGGGGGATTTCCCCTGCCCGTCTGCAGAAGGGTGAGCCAACCACCTTAGTAT  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CAGGGCGTGGGGGATTTCCCCTGCCCGTCTGCAGAAGGGTGAGCCAACCACCTTAGTAT  120

Qy   121 GTAGGCGGCGGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTAGGCGGCGGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180

Qy   181 CCTGATGGGCGTTCATGGGTTCGGTGGTGGTGGCGCCTTTAGGCAGCCTCCACGCCCACCA  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCTGATGGGCGTTCATGGGTTCGGTGGTGGTGGCGCCTTTAGGCAGCCTCCACGCCCACCA  240

Qy   241 CCTCCCAGATAGAGCGGCGGCACTGTAGGGAAGACCGGGGACCGGTCACTACCAAGGACG   300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCCCAGATAGAGCGGCGGCACTGTAGGGAAGACCGGGGACCGGTCACTACCAAGGACG   300

Qy   301 CAGACCTCTTTTGAGTATCACGGCTCCGGAAGTAGTTGGGCAGCCCACCTATATGTGT   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGACCTCTTTTGAGTATCACGGCTCCGGAAGTAGTTGGGCAGCCCACCTATATGTGT   360

Qy   361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGGAGGGGGAT  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGGAGGGGGAT  420

Qy   421 CTGGGAGTCTCGTAGACCGTAGCACATGCCGTGTTATTTCTACTCAAACAAGTCCTGTACC  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTGGGAGTCTCGTAGACCGTAGCACATGCCGTGTTATTTCTACTCAAACAAGTCCTGTACC  480

Qy   481 TGCGCCCAGAACGCGCAAGAACAGCAGACGCAGGCTTCATATCCGTGTCCATTAAAAC  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TGCGCCCAGAACGCGCAAGAACAGCAGACGCAGGCTTCATATCCGTGTCCATTAAAAC  540

Qy   541 ATCTGTTGAAAGGGGACAACGAGCAAAGCGCAAAGTCAGCGGACATGCTCGGCCTCGTAA  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 ATCTGTTGAAAGGGGACAACGAGCAAAGCGCAAAGTCAGCGGACATGCTCGGCCTCGTAA  600

Qy   601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC  660

Qy   661 TCATGGTTGGGGACGCCAAGACCCTCGCCAATCTGGAATCCTTCTGA  720
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TCATGGTTGGGGACGCCAAGACCCTCGCCAATCTGGAATCCTTCTGGA  720

Qy   721 TTACCCTTTGGGGTGGATTGGTGTAGTTACAACTCACACACCTCCAGTAGGCCCGCTGGT  780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TTACCCTGTAGTGGGGGTGGATTGGTGTAGTTACAACTCACACACCTCCAGTAGGCCCGCTGGT  780

Qy   781 GGCAGGAGCGGTCGTTCGACCAGTCTGCCAGTAGTACGCTTGCTGGAGGATGGAGTCAA  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GGCAGGAGCGGTCGTTCGACCAGTCTGCCAGTAGTACGCTTGCTGGAGGATGGAGTCAA  840

Qy   841 CTGGGCTACTGGTTGGTTGGTTCGGTATGTCGGTATGTCTGCTATCTTTGGCCTG  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CTGGGCTACTGGTTGGTTGGTTCGGTATGTCGGTATGTCTGCTATCTTTGGCCTG  900

Qy   901 TCCCTAGTGGGGCGCGGGGGTCACTGACCCAGACACAAATACCACCAATCCTGACCAATTG  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCCTGTAGTGGGGCGCGGGGGTCACTGACCCAGACACAAATACCACCAATCCTGACCAATTG  960

Qy   961 CTGCCAGCGTAATCAGGTTATCATTAGTGTTCTCCTTCCACTTGCCTACACGAGCCTGGTTG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CTGCCAGCGTAATCAGGTTATCATTAGTGTTCTCCTTCCACTTGCCTACACGAGCCTGGTTG 1020
```

```
Qy  1021  TGTGATCTGCGGACGAGTGCTGGGTTCCGCGCCAATCGTACATCCACACCCTTCCAA  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  TGTGATCTGCGGACGAGTGCTGGGTTCCGCGCCAATCGTACATCCACACCCTTCCAA  1080

Qy  1081  TTGGACTGGCACGGACTCCTTCTTGGCTGACCACATTGATTTGTTATGGGGCTCTTGT  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  TTGGACTGGCACGGACTCCTTCTTGGCTGACCACATTGATTTGTTATGGGGCTCTTGT  1140

Qy  1141  GACCTGTGAGCGCCCTTGACATTGGTGAGTTGTGTGGCGTGTATTAGTCGGTGACTG  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  GACCTGTGAGCGCCCTTGACATTGGTGAGTTGTGTGGCGTGTATTAGTCGGTGACTG  1200

Qy  1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGA  1260

Qy  1261  AGTGCCCACTGAATAGATCCTGGGTTCCTAGGGTTATCGGTGGATGGCCGGCAAGGT  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  AGTGCCCACTGAATAGATCCTGGGTTCCTAGGGTTATCGGTGGATGGCCGGCAAGGT  1320

Qy  1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAGTACCATACGCTATTGCGACTAT  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAGTACCATACGCTATTGCGACTAT  1380

Qy  1381  GTTTAGCAGTGTACACTACCTGGCGTGGTTGGCCGTCTGATCTACTACTGCCTCTCGGGGCAA  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  GTTTAGCAGTGTACACTACCTGGCGTGGTTGGCCGTCTGATCTACTACTGCCTCTCGGGGCAA  1440

Qy  1441  GTGGTATCAGTTGCTCCTAGCGCTTATGCTTACATAGAAGCGACCTCTGAAACCCTAT  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  GTGGTATCAGTTGCTCCTAGCGCTTATGCTTACATAGAAGCGACCTCTGAAACCCTAT  1500

Qy  1501  CAGGGTGCCCACTGAAGCTCAATAGCTGAGTTTTGCTCCGCCTTTGATGATACCATGTCC  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  CAGGGTGCCCACTGAAGCTCAATAGCTGAGTTTTGCTCCGCCTTTGATGATACCATGTCC  1560

Qy  1561  TTGCCACTCTATTTGAGTGAGAATGTCAGAGTCATTGTACAGTCCAAAGTGGAC  1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  TTGCCACTCTATTTGAGTGAGAATGTCAGAGTCATTGTACAGTCCAAAGTGGAC  1620

Qy  1621  CAGGCCTATCACTCTAGAGTATAAACTTCCATATCTTGGTACCCCTATACAATCCCTGG  1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  CAGGCCTATCACTCTAGAGTATAAACTTCCATATCTTGGTACCCCTATACAATCCCTGG  1680

Qy  1681  TGCGAGGGGGATGTATGGTTAAATTCAAAAATAACCATGGGGTTGCTGCCGTATTCGCAA  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  TGCGAGGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA  1740

Qy  1741  TGTGCCATCGTACTGACTACTGGGCACTATGGGCACTATGGGCAGTGTGGAACGCGACACTTA  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  TGTGCCATCGTACTGACTACTGGGCACTATGGGCACTATGGGCAGTGTGGAACGCGACACTTA  1800

Qy  1801  CGAAGCATGCGGTGTAACACCGTGTAACGGCACATGGCTCAACGGCTCAGCCCTGAA  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  CGAAGCATGCGGTGTAACACCGTGTAACGGCACATGGCTCAACGGCTCAGCCCTGAA  1860

Qy  1861  ATTGGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC  1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  ATTGGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC  1920

Qy  1921  AGGCCATTTGTATTTGTAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGAA  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  AGGCCATTTGTATTTGTAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGAA  1980

Qy  1981  TTCCACTCTCCTACCACCGGAGAGGTGGGCGTAGGTTGGCCCGGTACCCCACCTGTGGTACG  2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  TTCCACTCTCCTACCACCGGAGAGGTGGGCGTAGGTTGGCCCGGTACCCCACCTGTGGTACG  2040

Qy  2041  TGGTTCTTGGTTACAGGTTTTACAGTGATGTGAAAGAGTGAAAGACCTAGCCACCAGG  2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  TGGTTCTTGGTTACAGGTTTTACAGTGATGTGAAAGAGTGAAAGACCTAGCCACCAGG  2100
```

```
Qy  2101  ATTGATCACCAAAGACAAAGCTGGAAAAATATCAGGTCTTATATTCCGCCACGGGTGC  2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  ATTGATCACCAAAGACAAAGCTGGAAAAATATCAGGTCTTATATTCCGCCACGGGTGC  2160

Qy  2161  TTTGTCTCTTACGGGAGTGTACCACCAAGGCCGGTTGGTGTCAATTCGTGTGGGTTGTGTGG  2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TTTGTCTCTTACGGGAGTGTACCACCAAGGCCGGTTGGTGCTAATTCGTGTGGGTTGTGTGG  2220

Qy  2221  CAGCAAGTATCTTATTTGCCCTTGTCCCTTGTTTGGGGCGCTTC  2280
          ||||||||||||||||||||||||||||||||||||||||||||
Db  2221  CAGCAAGTATCTTATTTGCCCTTGTCCCTTGTTTGGGGCGCTTC  2280

Qy  2281  TGGTTACCCTTGCGTCCGTGCTCCATCCCAGTCGTATCTCGAAGCTGGCTGGGGATGT  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TGGTTACCCTTGCGTCCGTGCTCCATCCCAGTCGTATCTCGAAGCTGGCTGGGGATGT  2340

Qy  2341  TTTGTCTAAAGCTCAAGTAGCTAGCTCCTTTTGCTTGATTTCTTGATTTCTTCTCATCGTGTCTATCCG  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  TTTGTCTAAAGCTCAAGTAGCTAGCTCCTTTTGCTTGATTTCTTGATTTCTTCTCATCGTGTCTATCTCG  2400

Qy  2401  CTGCAGGCTACGTATGCGTCGCCCACCAGATTATGACTGGTGGGTGGTGGACTGCT  2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  CTGCAGGCTACGTATGCGTCGCCCACCAGATTATGACTGGTGGGTGGTGGACTGCT  2460

Qy  2461  AACTTTCTTTGTGCAGCAGCTGCTGCCCAACCGCGTAACCGTGCTCACCGGAAACCCAT  2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  AACTTTCTTTGTGCAGCAGCTGCTGCCCAACCGCGTAACCGTGCTCACCGGAAACCCAT  2520

Qy  2521  AGTGGCAGGGTTAGTTTTGTGGGCCGGCCGTAACCGGTGGTCACCGCATAGCTCGCTTGT  2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  AGTGGCAGGGTTAGTTTTGTGGGCCGGCCGTAACCGGTGGTCACCGCATAGCTCGCTTGT  2580

Qy  2581  AGGTCCTTGGGCCTGGTAGCGCGTTTAACCCTCTTGATTTGGTTACGCCTGCTTCAGC  2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  AGGTCCTTGGGCCTGGTAGCGCGTTTAACCCTCTTGATTTGGTTACGCCTGCTTCAGC  2640

Qy  2641  TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT  2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  TTTTGACACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT  2700

Qy  2701  GTCTCGTTTGGCTTCTTGCTCACTTGTTGCTCCGCTCGCTGTCGTGCTTTAGTTAACTCCTATCT  2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  GTCTCGTTTGGCTTCTTGCTCACTTGTTGCTCCGCTCGCTGTCGTGCTTTAGTTAACTCCTATCT  2760

Qy  2761  TTGGCAACGTTGGAGAATTGTTTCCCGGTGCTGCCGTATTCGCAA  2820
          |||||||||||||||||||||||||||||||||||||||||||||
Db  2761  TTGGCAACGTTGGAGAATTGTTTCCCGGTGCTGCCGTATTCGCAA  2820

Qy  2821  TGTGCTCGGTTGTTGTTTCCCCGGTGCGCACATATGACGCGGCTGGTGACTTTCTCGTGTGTGTCA  2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  TGTGCTCGGTTGTTGTTTCCCCGGTGCGCACATATGACGCGGCTGGTGACTTTCTCGTGTGTGTCA  2880

Qy  2881  CGTAGCTCTTCATGTGTTCAGTCGTTAACATCAGTGCAGCGCAGCATCGTCTTGTTCCT  2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  CGTAGCTCTTCATGTGTTCAGTCGTTAACATCAGTGCAGCGCAGCATCGTCTTGTTCCT  2940

Qy  2941  TAGGGCCCATAGAAGTGTTGGTGCGTCTCGGAAAGTGTCCGGATGTC  3000
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  2941  TAGGGCCCATAGAAGTGTTGGTGCGTCTCGGAAAGTGTCCGGATGTC  3000

Qy  3001  TCTTAAGTTTTTCCCTCTAGTGTTTAGTTTTCTATAAGCACTTGCA  3060
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  TCTTAAGTTTTTCCCTCTAGTGTTTAGTTTTCTATAAGCACTTGCA  3060

Qy  3061  TGGTGATGTCTTGCCTAATGATTTGCCTCGAAATGATGCCATTTTTCCC  3120
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061  TGGTGATGTCTTGCCTAATGATTTGCCTCGAAATTGCAATTTTTCCC  3120

Qy  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACGCTTGGGCGTGTGGGGACAC  3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACGCTTGGGCGTGTGGGGACAC  3180

Qy  3181  GGTTGATGGTTGCCCGGTTGTTGCCCGTCTCGGCGACCTTGTTTCGGCAGGGTTGGCTAT  3240
```

Db |||||||||||||||||||||||||||||||||||||||||||| GGTTGATGGTTGCCCGTGTTGCGGCGTCCGGCCACCTTGTTTCGGAGGGTTAGCTAT 3240

Qy 3241 GCCGCCAGATGGTGGGCCATTACCGCACCGTTTACGCTGCGCAGTGTCTTCTGAACGTGG 3300
Db 3241 GCCGCCAGATGGTGGGCCATTACCGCACCGTTTACGCTGCGCAGTGTCTTCTGAACGTGG 3300

Qy 3301 CACGCTGTCAGCGATGGCAGTGGTCATGACTGTATAGACACCCCGACTTGGACTGGAAC 3360
Db 3301 CACGCTGTCAGCGATGGCAGTGGTCATGACTGTATAGACACCCCGACTTGGACTGGAAC 3360

Qy 3361 TATCTTCAGATTAGGGATCTCTGGCCACTAGCTACAACAGGGATTGTTTGTGACAACGTGTT 3420
Db 3361 TATCTTCAGATTAGGGATCTCTGGCCACTAGCTACAACAGGGATTGTTTGTGACAACGTGTT 3420

Qy 3421 GTATACTGCTCACCATGGCAGCAAGGGGCGCCGGTTGCTCATCCCACAGGCTCTATACA 3480
Db 3421 GTATACTGCTCACCATGGCAGCAAGGGGCGCCGGTTGCTCATCCCACAGGCTCCATACA 3480

Qy 3481 CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG 3540
Db 3481 CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG 3540

Qy 3541 GTCCCTTACTCGGTGCTCTTGCGGGAGACCAAGGGGTATCTGGTAACGACTGGGGTC 3600
Db 3541 GTCCCTTACTCGGTGCTCTTGCGGGAGACCAAGGGGTATCTGGTAACGACTGGGGTC 3600

Qy 3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTATTGGTGTGTGCGGGGCCCTTCCCAT 3660
Db 3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTATTGGTGTGTGCGGGGCCCTTCCCAT 3660

Qy 3661 GGCTGTTGCCAAGGGTTCTTCAGTGTGCCCGATTCTGTGCTCCTCCGGGCATGTTATTGG 3720
Db 3661 GGCTGTTGCCAAGGGTTCTTCAGTGTGCCCGATTCTGTGCTCCTCCGGGCATGTTATTGG 3720

Qy 3721 GATGTTCACCGGTGCTAGAAATTCTGGCGGTTCAGTCAGTTAGGCCGTT 3780
Db 3721 GATGTTCACCGGTGCTAGAAATTCTGGCGGTTCAGTCAGTTAGGCCGTT 3780

Qy 3781 GGTGTGTGCTGGATACCAGTACACAGACACATGCCACTCTTGATACAAAACCTAC 3840
Db 3781 GGTGTGTGCTGGATACCAGTACACAGACACATGCCACTCTTGATACAAAACCTAC 3840

Qy 3841 TGTGCCTAACGACTATTCAGTGCAAATTTAATTGCCCCCACTGGAGCAGCGGCAAGTCAAC 3900
Db 3841 TGTGCCTAACGACTATTCAGTGCAAATTTAATTGCCCCCACTGGAGCAGCGGCAAGTCAAC 3900

Qy 3901 CAAATTACCACTTTCTTACAGGAGGAAGTAGTGAGGTCTTGGTCCTAAATCCCAGTGT 3960
Db 3901 CAAATTACCACTTTCTTACAGGAGGAAGTAGTGAGGTCTTGGTCCTAAATCCCAGTGT 3960

Qy 3961 GGCTACAACAGCATCAATGCCAAAGTCACGCGACGTACGGGGCTTCACTCAATTG 4020
Db 3961 GGCTACAACAGCATCAATGCCAAAGTCACGCGACGTACGGGGCTTCACTCAATTG 4020

Qy 4021 CTATTTTAATGGCAAATGTACCAACACAGGGGCTTCACTTACGGGCACACATATGGCAT 4080
Db 4021 CTATTTTAATGGCAAATGTACCAACACAGGGGCTTCACTTACGGGCACACATATGGCAT 4080

Qy 4081 GTACCTGACCGGAGCATGTTCCCGGAACTATGATGATGTCACATTGTGACGAATGGCCATGC 4140
Db 4081 GTACCTGACCGGAGCATGTTCCCGGAACTATGATGATGTCACATTGTGACGAATGGCCATGC 4140

Qy 4141 TACCTGGTGCTTCTTGCCACGGCGTGTTGGGCATTGGAAAGGTTCAACCGAAGTTCCATCCAAAAA 4200
Db 4141 TACCTGGTGCTTCTTGCCACGGCGTGTTGGGCATTGGAAAGGTTCAACCGAAGTTCCATCCAAAAA 4200

Qy 4201 TGTTAGGCTAGTGTTCTTGCCACGGCGTACCCCCTGAGTAATCCCCTGAGTAATCCCCTGAGTAATGCCTACACCACACATGC 4260
Db 4201 TGTTAGGCTAGTGTTCTTGCCACGGCGTACCCCCTGAGTAATCCCCTGAGTAATCCCCTGAGTAATGCCTACACCACACATGC 4260

Qy 4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTCATGCGGAAAAAAGAT 4320
Db 4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTCATGCGGAAAAAACTG 4320

Db   CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTCATGCGGAAAAAAGAT 4320

Qy 4321 TAAGGAGGAAAAATCTGAAGAAAAGGAGGAGGACACCTTATCTTGAGGCTACCACACTG 4380
Db 4321 TAAGGAGGAAAAATCTGAAGAAAAGGAGGAGGACACCTTATCTTGAGGCTACCAAAAACACTG 4380

Qy 4381 TGATGAGCTTGCTAACGAGTTAGCTGCAAAGGAATAACAGCTGTCTCTTACTATAGGGG 4440
Db 4381 TGATGAGCTTGCTAACGAGTTAGCTGCAAAGGAATAACAGCTGTCTCTTACTATAGGGG 4440

Qy 4441 ATGTGACATCTCAAAAATCCCTGAGGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500
Db 4441 ATGTGACATCTCAAAAATCCCTGAGGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500

Qy 4501 TACAGGGTACACTGGTGACTTTGATTCCGTGATGACTGCTCAGCCTCATGGTAGAAGGCAC 4560
Db 4501 TACAGGGTACACTGGTGACTTTGATTCCGTGATGACTGCTCAGCCTCATGGTAGAAGGCAC 4560

Qy 4561 ATGCCATGTTGACCTTGACCTACTTTCACCATGGCGTGTCGTGTGTGCGGGGTTTCAGC 4620
Db 4561 ATGCCATGTTGACCTTGACCTACTTTCACCATGGCGTGTCGTGTGTGCGGGGTCTCAGC 4620

Qy 4621 AATAGTTAAAGGCCAGCCACAGCCGCGTAGGGGTAGGGGCCGCACAGCCTGGCATATACTACTA 4680
Db 4621 AATAGTTAAAGGCCAGCCACAGCCGCGTAGGGGTAGGGGCCGCACAGCCTGGCATATACTACTA 4680

Qy 4681 TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740
Db 4681 TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740

Qy 4741 CGACGCAGCCAAGGCATGTATGGTTGTCATCAACAGAAGCTCAAACTATTCTGGACAC 4800
Db 4741 CGACGCAGCCAAGGCATGTATGGTTGTCATCAACAGAAGCTCAAACTATTCTGGACAC 4800

Qy 4801 CTATCGCACCAACCTGGTTACCTGCAGATAGAGCAAATTTGGACGACGTGGGCGTGATCT 4860
Db 4801 CTATCGCACCAACCTGGTTACCTGCAGATAGAGCAAATTTGGACGACGTGGGCGTGATCT 4860

Qy 4861 CTTTTCTATGTGCAACCCGAACCTTCATTGTTCAATACTGTCAATACTGCTGACAA 4920
Db 4861 CTTTTCTATGTGCAACCCGAACCTTCATTGTTCAATACTGTCAATACTGCTGACAA 4920

Qy 4921 TTATGTTTTGTTGACTGCAGCGCAGCCAACTGTGTCATCAGTATGGCTATGCTGCTCC 4980
Db 4921 TTATGTTTTGTTGACTGCAGCGCAGCCAACTGTGTCATCAGTATGGCTATGCTGCTCC 4980

Qy 4981 CAATGACGGCACCACGTGCAGGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG 5040
Db 4981 CAATGACGGCACCACGTGCAGGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG 5040

Qy 5041 GCGCTTGGACGGGCGGCCTGACGCCGTTTTGGCGCCGCGAGGTGACCAGATACCA 5100
Db 5041 GCGCTTGGACGGGCGGCCTGACGCCGTTTTGGCGCCGCGAGGTGACCAGATACCA 5100

Qy 5101 AATGTGCTTCACTGAAGTCAATACTTCGGCGTTGGCGTTGGCGGCGTTGCTGGAGT 5160
Db 5101 AATGTGCTTCACTGAAGTCAATACTTCGGCGTTGGCGTTGGCGGCGTTGCTGGAGT 5160

Qy 5161 GGCTATGGCTTATCTAGCCATTAGCACTTTGACACTTTTGGCGCTCGCCCAGGGTTGTTCGCGGTC 5220
Db 5161 GGCTATGGCTTATCTAGCCATTAGCACTTTGACACTTTTGGCGCTCGCCCAGGGTTGTTCGCGGTC 5220

Qy 5221 TATTACATCAGTCCTACCGGTGCTACTGTGTTGACGAAGAAGAAATCGT 5280
Db 5221 TATTACATCAGTCCTACCGGTGCTACTGTGTTGACGAAGAAGAAATCGT 5280

Qy 5281 GGAGGAGTGTGCATCATTCCCCTTGGAGGCCATGGTTGCTGCTAATCGCGATAAGCTGAA 5340
Db 5281 GGAGGAGTGTGCATCATTCCCCTTGGAGGCCATGGTTGCTGCTAATCGCGATAAGCTGAA 5340

Qy 5341 GAGTACAATCACCACAACTAGTCCTTTCATTGGAAAACCGCCCTTGAAAAACTTAACAC 5400
Db 5341 GAGTACAATCACCACAACTAGTCCTTTCATTGGAAAACCGCCCTTGAAAAACTTAACAC 5400

Left column:

```
Qy  5401  CTTTCTTGGGCCTCATGCGCTCATAGCAGTCACAATCCTTGCTATCATAGAGTATTGCTGTGGTTTAGT  5460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401  CTTTCTTGGGCCTCATGCGCTCATAGCAGTCACAATCCTTGCTATCATAGAGTATTGCTGTGGCTTAGT  5460

Qy  5461  CACTTTACCTGACAATCCCTTTGCATCATGCCGTGTTTGCTTTCATTGCGGGTATTACTAC  5520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461  CACTTTACCTGACAATCCCTTTGCATCATGCCGTGTTTGCTTTCATTGCGGGTATTACTAC  5520

Qy  5521  CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTGGAGGCGCAATTGCGTCCAA  5580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521  CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTGGAGGCGCAATTGCGTCCAA  5580

Qy  5581  GCTTACAGACGCTAGAGGCGCACTGGCCGTTCATGATGGCGCTCGTGGGGAACAGCTCT  5640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581  GCTTACAGACGCTAGAGGCGCACTGGCCGTTCATGATGGCGCTCGTGGGGAACAGCTCT  5640

Qy  5641  TGGTACATGGACATCGGTGGGTTTGCTTTGCTTTCATTGCGGCGGTATGCTGCCGCGTC  5700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641  TGGTACATGGACATCGGTGGGTTTGCTTTGCTTTCATTGCGGCGGTATGCTGCCGCGTC  5700

Qy  5701  ATCCACTGCTTGCTGACATTTAAATGCTTGATGGGTGAGTGGCCACTATGGATCAGCT  5760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701  ATCCACTGCTTGCTGACATTTAAATGCTTGATGGGTGAGTGGCCACTATGGATCAGCT  5760

Qy  5761  TGCTGGTTTAGTCTACTCCGCGTTCAATCCGGCCAGGGGCCAGGAGTTGTGCGGTCTTGTCAGC  5820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761  TGCTGGTTTAGTCTACTCCGCGTTCAATCCGGCCAGGGGCCAGGAGTTGTGCGGTCTTGTCAGC  5820

Qy  5821  TTGTGCAATGTTTGCTTTGACAACAGCAGGGCCAGATCACTGGCCCAACAGACTTCTTAC  5880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821  TTGTGCAATGTTTGCTTTGACAACAGCAGGGCCAGATCACTGGCCCAACAGACTTCTTAC  5880

Qy  5881  TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCCGTGACATCCG  5940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881  TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCCGTGACATCCG  5940

Qy  5941  CAGGAAGATACTGGGCATTCTCTGGAGCATCATCACCCCCCTGGAGTGTCATATCAGCTTGCAT  6000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941  CAGGAAGATACTGGGCATTCTCTGGAGCATCATCACCCCCCTGGAGTGTCATATCAGCTTGCAT  6000

Qy  6001  CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGGCCTCATTGCGGGCCTCATTGCGGGCCC  6060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001  CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGGCCTCATTGCGGGCCTCATTGCGGGCCC  6060

Qy  6061  TTGGCAGTATGTGCAATTTCTTGTGATTGCTTTAAATGTCCTTAAAGCTGGAGTTCA  6120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061  TTGGCAGTATGTGCAATTTCTTGTGATTGCTTTAAATGTCCTTAAAGCTGGAGTTCA  6120

Qy  6121  GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAAGAGGGGTACAAGGGCCC  6180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121  GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAAGAGGGGTACAAGGGCCC  6180

Qy  6181  CTGGATTGGATCAGGTATGCTCCAAGCACGCGTGCTGAACTCATCTTTTC  6240
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181  CTGGATTGGATCAGGTATGCTCCAAGCACGCGTGCTGAACTCATCTTTTC  6240

Qy  6241  TGTTGAGAATGGTTTGCAAATGTTTGCAATTGCTTGTTCAAATTACTGGAG  6300
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241  TGTTGAGAATGGTTTGCAAATGTTTGCAATTGCTTGTTCAAATTACTGGAG  6300

Qy  6301  AGGGGCTGTTCCAGTCAACGCTAGGCGTGTGTGGGTCGGTAGGACCGGACCCAACTGATTG  6360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301  AGGGGCTGTTCCAGTCAACGCTAGGCGTGTGTGGGTCGGTAGGACCGGACCCAACTGATTG  6360

Qy  6361  GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA  6420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361  GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA  6420

Qy  6421  TCACATTTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCCAGTGCCCCCAAC  6480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421  TCACATTTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCCAGTGCCCCCAAC  6480
```

Right column:

```
Qy  6481  CTTGAGAGCTGCAGTGGCCGTGACGGCGTGCAGGTTCAGTGTGTTATCTAGGTGAGCCCAA  6540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481  CTTGAGAGCTGCAGTGGCCGTGACGGCGTGCAGGTTCAGTGTGTTATCTAGGTGAGCCCAA  6540

Qy  6541  AACTCCTTGGACGACATCATCGCTTGCTCGTCTGCTGCTCCGACGGTAAGGTAAAACTGTTAA  6600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541  AACTCCTTGGACGACATCATCGCTTGCTCGTCTGCTGCTCCGACGGTAAGGTAAAACTGTTAA  6600

Qy  6601  GCTTCCCTTCCGGTTGACGGTCACACACCTGGTGTGGCATGGCAACTTAATTGCGTGA  6660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601  GCTTCCCTTCCGGTTGACGGTCACACACCTGGTGTGGCATGGCAACTTAATTGCGTGA  6660

Qy  6661  TGCACTTGAGACAAATGACTGTAATTCCACAAACACTCCTAGTGATGAGCCGCGAGT  6720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661  TGACTTGAGACAAATGACTGTAATTCCATAAACACTCCTAGTGATGAAGCCGCGAGT  6720

Qy  6721  GTCCGCTCCTTGTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTGAGGCAATTTC  6780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721  GTCCGCTCCTTGTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTGAGGCAATTTC  6780

Qy  6781  AGCTGGCGTTGACCACCAAACTGCCAGCCCCCTCCATCGAAGAGGTAGTGGTAAGAAA  6840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781  AGCTGGCGTTGACCACCAAACTGCCAGCCCCCTCCATCGAAGAGGTAGTGGTAAGAAA  6840

Qy  6841  GCGCCAGTTCCGGGCAAGAACTGGTTCGGCTTACCCTTGCCTCCCCCCTCCGAGATCCGTCCC  6900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841  GCGCCAGTTCCGGGCAAGAACTGGTTCGGCTTACCCTTGCCTCCCCCCTCCGAGATCCGTCCC  6900

Qy  6901  AGGAGTGTCATGTCCTGCAAGCTGCTGCGCAACGAGTGACCCGTGGCCATGCGAAAGCTGATGA  6960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901  AGGAGTGTCATGTCCTGCAAGCTGCTGCGCAACGAGTGACCCGTGGCCATGCGAAAGCTGATGA  6960

Qy  6961  CCCTCCTTCTTCACCACCTTTCACTGCAATTGGCATGTGCAATGACCGGGAGCGGCCCTGATGA  7020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961  CCCTTCTTCTTCACCACCTTTCACTGCAATTGGCATGTGCAATGACCGGGAGCGGCCCTGATGA  7020

Qy  7021  GTGTAACCCTTCACTGCAATTGGATGTGCAATGGTCAGACGGTCTGAATGGTCGGAAAGAGTTGGTCGAC  7080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021  GTGTAACCCTTCACTGCAATTGGATGTGCAATGGTCAGACGGTCTGAATGGTCGGAAAGAGTTGGTCAAC  7080

Qy  7081  TTTACCCAGTAGTCACGCTTCCCCAAAAAGGAGGTCTCTGAATGGTCAGACCGGGTACCCGTGATGGTCGAC  7140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081  TTTACCCAGTAGTCACGCTTCCCCAAAAAGGAGGTCTCTGAATGGTCAGACCGGGTACCCGTGATGGTCAAC  7140

Qy  7141  GGCTACAACCGCTTCCAGTCACGCTACTGTGGCCCCCGTACCCGTAAGATACGGGAAAGGA  7200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141  GACTACAACCGCTTCCAGTCACGCTACTGTGGCCCCCGTACCCGTAAGATACGGGAAAGGA  7200

Qy  7201  TTCCACTCAGTGACCGCCGGGATGCGGAGCTTAGAAGCAGCGGCCAAACGGGGTACAACGGGGCCC  7260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201  TTCCACTCAGTGACCGCCGGGATGCGGAGCTTAGAAAACAAAAGTCACTATTAATAGACAACCTCT  7260

Qy  7261  TTCGTGCAGCATGAGCTAGCCGACCGTGATTGGCCTAAGGACAAGCTTCAAAAAGTTGT  7320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261  TTCGTGCAGCATGAGCTAGCCGACCGTGATTGGCCTAAGGACAAGCTTCAAAAAGTTGT  7320

Qy  7321  TCTGTCTGCAACTCGGGGCCATCAGTGGTTTCCTCAAACAAAGATCATTGGTGTATGT  7380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7321  TCTGTCTGCAACTCGGGGCCATCAGTGGTTTCCTCAAACAAAGATCATTGGTATGT  7380

Qy  7381  GACTGAGCCGCCGGGATGCGGAGCTTAGAAAAACAAAAGTCACTATTAATAGACAACCTCT  7440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381  GACTGAGCCGCCGGGATGCGGAGCTTAGAAAAACAAAAGTCACTATTAATAGACAACCTCT  7440

Qy  7441  GTTCCCCCCATCATACCACACAACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT  7500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441  GTTCCCCCCATCATACCACACAACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT  7500

Qy  7501  CGGTGTCATGTGGACTATGATGAATGATGATAGCAGCTGAAGTTCACGCGCCCCTTCTCAAGTCTGCTAAGTC  7560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7501  CGGTGTCATGTGGACTATGATGAATGATGATAGCAGCTGAAGTTCACGCGCCCCTTCTCAAGTCTGCTAAGTC  7560

Qy  7561  CCACATCACTGGCCTTCGGGGCACTGATGTTCGTTCTGGGACGCCCGAAGGCGTGTTCT  7620
Db  7561  
```

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
7561 CCACATCACTGGCCTTCGGGGCACTGATGTTCGTTCTGGAGCAGCCCGCAAGGCTGTTCT 7620

Qy   GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680

Qy   AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAAACCAACACAAGAAACCCCCAAG 7740
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAAACCAACACAAGAAACCCCCAAG 7740

Qy   GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAGATGTACTACGGTCAGGT 7800
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
7741 GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAGATGTACTACGGTCAGGT 7800

Qy   TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTACGGGTTTGTAGATCCACGTAC 7860
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTACGGGTTTGTAGATCCACGTAC 7860

Qy   CCGTGTCAAGCGTCTGTGTCGATGTGTCAGTCGGAGCCACATGCGATAC 7920
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
7861 CCGTGTCAAGCGTCTGTGTCGATGTGTCAGTCGGAGCCACATGCGATAC 7920

Qy   AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTTGGAGACAGACATCTACTC 7980
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
7921 AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTTGGAGACAGACATCTACTC 7980

Qy   AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCATTCACACCATTGCGAGGCAGTTATA 8040
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCATTCACACCATTGCGAGGCAGTTATA 8040

Qy   CGCTGGAGGACCGATGATCGCTTATGATGGCCGAGATCGGATATCGTAGGTGTGTAGGTC 8100
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8041 CGCTGGAGGACCGATGATCGCTTATGATGGCCGAGATCGGATATCGTAGGTGTGTAGGTC 8100

Qy   TTCCGGGCGTCTATACTACCTCAAGTTCCAACAGTTGACCTGCTGGCTGAAGTAAATGC 8160
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8101 TTCCGGGCGTCTATACTACCTCAAGTTCCAACAGTTGACCTGCTGGCTGAAGTAAATGC 8160

Qy   TGCAGCCGAACAGGCTGGCATGGAAGAACCCTCGCTTCCTTATTGCGGCGATGATTGCAC 8220
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8161 TGCAGCCGAACAGGCTGGCATGGAAGAACCCTCGCTTCCTTATTGCGGCGATGATTGCAC 8220

Qy   CGTAAATTGGAAGAGCGGCGGCAGATGCAGACAAGCAATGCGTGTCTTTGCTAG 8280
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8221 CGTAAATTGGAAGAGCGGCGGCAGATGCAGACAAGCAATGCGTGTCTTTGCTAG 8280

Qy   CTGGATGAAGGTGATGGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8281 CTGGATGAAGGTGATGGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340

Qy   AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400

Qy   CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCCAGGTGTCTCTGCCGAGGGTCTGGG 8460
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCCAGGTGTCTCTGCCGAGGGTCTGGG 8460

Qy   ATACAACCCAGTGCTGCGTGATGGGTTCGGGTATCTAATACATCACTACCCCATGTTTGTGGGT 8520
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8461 ATACAACCCAGTGCTGCGTGATGGGTTCGGGTATCTAATACATCACTACCCCATGTTTGTGGGT 8520

Qy   TAGCCGTGTGTTGGCGTGTCCATTTCATGGAGCAGAGATGCTCTTTGAGGACAAACTTCCCGA 8580
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8521 TAGCCGTGTGTTGGCGTGTCCATTTCATGGAGCAGAGATGCTCTTTGAGGACAAACTTCCCGA 8580

Qy   GACTGTGACCTTTGACTGTGTATGGGAAAAATTATACGGTGCCGTGTAGAAGATCTGCCCAG 8640
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8581 GACTGTGACCTTTGACTGTGTATGGGAAAAATTATACGGTGCCGTGTAGAAGATCTGCCCAG 8640

Qy   CATCATTGCTCGGTGTGCACGGTATTGAGGCTTTCTCCGGTGGTGCGCTACACCAACGCTGA 8700

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8641 CATCATTGCTCGGTGTGCACGGTATTGAGGCTTTCTCGGTGGTGCGCTACACCAACGCTGA 8700

Qy   GATCCTCAGAGTTTCCCAATCACTCAACACAGACACATGACACATGACCACATGCCCCCCTGCGAGCCTGGCG 8760
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8701 GATCCTCAGAGTTTCCCAATCACTCAACACAGACACATGACACATGACCACATGCCCCCCTGCGAGCCTGGCG 8760

Qy   AAAGAAAGCCAGGGCGGTCTCGCGAACCGGCGGTGGCGGAGCACACGCAAATT 8820
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8761 AAAGAAAGCCAGGGCGGTCTCGCGAACCGGCGGTGGCGGAGCACACGCAAATT 8820

Qy   GGCTCGCTTCCTTCTCTGCATGCTACATCTAGACCTCTACCAGATTGGATAAGACGAG 8880
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8821 GGCTCGCTTCCTTCTCTGCATGCTACATCTAGACCTCTACCAGATTGGATAAGACGAG 8880

Qy   CGTGGCTCGGTACCACCTTCAATTATGTGATGTTTACTCCCGGAGGGGGATGTGTT 8940
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8881 CGTGGCTCGGTACCACCTTCAATTATGTGATGTTTACTCCCGGAGGGGGATGTGTT 8940

Qy   TATTACACCAGAGAAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTTT 9000
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8941 TGTTACACCAGAGAAGATTGCAGAAGTTCTGTGAAGTATTTGGCTGTCATTGTTTT 9000

Qy   TGCCCTAGGGCTCATTGTGTTGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
9001 TGCCCTAGGGCTCATTGTGTTGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060

Qy   CTAACAG---TTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGGAGCAGCGGCAACAGGGGAGACCCC 9116
Db ||||||| |||||||||||||||||||||||||||||||||||||||||||||||
9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGGAGCAGCGGCAACAGGGGAGACCCC 9120

Qy   GGGCTTAACGACCCCGC 9133
Db |||||||||||||||||
9121 GGGCTTAACGACCCCGC 9137

RESULT 8
AAA55379
ID   AAA55379 standard; DNA; 9143 BP.
XX
AC   AAA55379;
XX
DT   06-AUG-2003   (revised)
DT   30-AUG-2000   (first entry)
XX
DE   Hepatitis GB virus nucleotide sequence SEQ ID NO:393.
XX
KW   Hepatitis GB virus; HGBV; diagnosis; therapeutic; immunogenic; infection;
KW   detection; characterisation; hepatitis; ds.
XX
OS   Hepatitis GB virus.
XX
PN   US6051374-A.
XX
PD   18-APR-2000.
XX
PF   07-JUN-1995;   95US-00488445.
XX
PR   14-FEB-1994;   94US-00196030.
PR   13-MAY-1994;   94US-00242654.
PR   29-JUL-1994;   94US-00283314.
PR   23-NOV-1994;   94US-00344185.
PR   23-NOV-1994;   94US-00344190.
PR   30-JAN-1995;   95US-00377557.
XX
PA   (ABBO) ABBOTT LAB.
XX
PI   Dawson GJ, Leary TP, Muerhoff AS, Pilot-Matias TJ, Buijk SL;
PI   Mushahwar IK, Simons JN, Desai SM, Erker JC, Schlauder GG;
XX
DR   WPI; 2000-338307/29.
XX
PT   Detecting target hepatitis GB virus nucleic acid in a test sample
PT   suspected of containing HGBV comprises reacting the test sample the HGBV

PT   polynucleotide probe and detecting the complex that contains target HGBV.
XX
PS   Example 9; Col 419-438; 369pp; English.
XX
CC   The present invention describe a method for detecting target hepatitis GB
CC   virus (HGBV) nucleic acid (THN) in a test sample (T) suspected of
CC   containing HGBV. The method involves reacting (T) with a HGBV
CC   polynucleotide probe (I) containing 15 contiguous nucleotides, and which
CC   selectively hybridises to the HGBV genome or its full complement, and
CC   detecting the complex that contains THN, indicating the presence of
CC   target HGBV. The method is used for detecting target HGBV nucleic acid in
CC   the test sample suspected of containing HGBV and for characterisation of
CC   newly ascertained etiological agent of non-A, non-B, non-C, non-D and non
CC   -E hepatitis causing agents collectively termed as hepatitis GB virus.
CC   AAA55270 to AAA55489 and AAB08985 to AAB09480 represent nucleotide and
CC   protein sequences used in the exemplification of the present invention.
CC   (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ   Sequence 9143 BP; 2071 A; 2266 C; 2349 G; 2457 T; 0 U; 0 Other;

Query Match   96.4%;   Score 9059.8;   DB 3;   Length 9143;
Best Local Similarity   99.6%;   Pred. No. 0;
Matches 9096;   Conservative   0;   Mismatches   37;   Indels   4;   Gaps   1;

Qy   1 ACCAAAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGAGCACCCCCCTAG   60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCAAAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGAGCACCCCCCTAG   60

Qy   61 CAGGGGCTGGGGGGATTTCCCCTGCCCCTCGCGTGATGACAGCGCCAAGCTTGACTGGATGGC   120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CAGGGGCTGGGGGGATTTCCCCTGCCCCTCGCGTGATGACAGCGCCAAGCTTGACTGGATGGC   120

Qy   121 GTAGGCGGCGGGACTCATGACGCCTCGCGTGATGACAAGCGCCAAGCGCCAAGCTTGGATGGC   180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTAGGCGGCGGGACTCATGACGCCTCGCGTGATGACAAGCGCCAAGCGCCAAGCTTGGATGGC   180

Qy   181 CCTGATGGGCGTTCATGGGTTCGGTGGTGGTGGTGGCGCTTAGGCAGCCTCCACGCCCACCA   240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCTGATGGGCGTTCATGGGTTCGGTGGTGGTGGTGGCGCTTAGGCAGCCTCCACGCCCACCA   240

Qy   241 CCTCCCAGATAGAGCCGGCCACTGTAGGGAAGGAAGACCCGGGACCGGTCACTACCACCAAGGACG   300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCCCAGATAGAGCCGGCCACTGTAGGGAAGGAAGACCCGGGACCGGTCACTACCACCAAGGACG   300

Qy   301 CAGACCTCTTTTGAGTATCACGCCCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTGT   360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGACCTCTTTTGAGTATCACGCCCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTGT   360

Qy   361 TGGGATGGTTGGGGTTAGCCATCCATCATCCGTACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT   420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TGGGATGGTTGGGGTTAGCCATCCATCATCCGTACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT   420

Qy   421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTGTTATTTCTACTACTACAACAAGTCCTGTACC   480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTGTTATTTCTACTACTACAACAAGTCCTGTACC   480

Qy   481 TGCGCCCAGAACGCGCAAGAACAACGCAGCAGACGCGAGGCTTCATATCCTGTGTCATTAAAAC   540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TGCGCCCAGAACGCGCAAGAACAACGCAGCAGACGCGAGGCTTCATATCCTGTGTCATTAAAAC   540

Qy   541 ATCTGTTGAAAGGGACAACGCGCAAAGCCAAAGTCCAGCGCGATGCTCCGGCCTCGTAA   600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 ATCTGTTGAAAGGGACAACGCGCAAAGCCAAAGTCCAGCGCGATGCTCCGGCCTCGTAA   600

Qy   601 TTACAAAATTGCTGGTATCCATGATGCTTGCAGACATTGGCTAGGCACATTCGCTTGCCAGC   660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTACAAAATTGCTGGTATCCATGATGCTTGCAGACATTGGCTAGGCACATTCGCTTGCCAGC   660

Qy   661 TCATGGTTGGGGACGACCCAAGACCCTGCCAATCTCGCAATCTTGGAATCTTCGTCCTTCTGGA   720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TCATGGTTGGGGACGACCCAAGACCCTGCCAATCTCGCAATCTTGGAATCTTCGTCCTTCTGGA   720

Qy   721 TTACCCTTGGGGTGGATTGGTGATGTGATGCTGCCAGATAGTAGTTACAACTCACACACCCTGGT   780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TTACCCTTGGGGTGGATTGGTGATGTGATGCTGCCAGATAGTAGTTACAACTCACACACCCTGGT   780

Db   781 GGCAGGAGCGGCGTCGTTGCACCAGTAGTACGCTTGCTGCTGGAGGATGGAGTCAA   840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   781 GGCAGGAGCGGCGTCGTTGCACCAGTAGTACGCTTGCTGCTGGAGGATGGAGTCAA   840

Qy   841 CTGGGCTACTGGTTCGGTGTCCACCTTTGTGGTATGTGCTGCATCTTGGCCTG   900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CTGGGCTACTGGTTCGGTGTCCACCTTTGTGGTATGTGCTGCATCTTGGCCTG   900

Qy   901 TCCCTGTAGTGGGCGCGGGTCACTGACCAGACACAAATACCACAATTGACCAATTG   960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCCTGTAGTGGGCGCGGGTCACTGACCAGACACAAATACCACAATTGACCAATTG   960

Qy   961 CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCTCCACTTGCTACGAGCCTGGTTG   1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCTCCACTTGCTACACGAGCCTGGTTG   1020

Qy   1021 TGTGATCTGTGCGGACGAGTGCTTGGGGTTCCCGGACCGTACAATCCGTACACCTTCAA   1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TGTGATCTGTGCGGACGAGTGCTTGGGGTTCCCGGACCGTACAATCCGTACACCTTCAA   1080

Qy   1081 TTGGACTGGCACGGACTCCTTCTGGCTGACCACCATTGTTATTGGGGCTCTTGT   1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TTGGACTGGCACGGACTCCTTCTGGCTGACCACCATTGTTATTGGGGCTCTTGT   1140

Qy   1141 GACCTGTGACGCCCTTGACATTGCGAGTGTGTGGTCGGTGTATTAGTCGGTGACTG   1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTGTGACGCCCTTGACATTGCGAGTGTGTGGTCGGTGTATTAGTCGGTGACTG   1200

Qy   1201 GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTACCTGGA   1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTACCTGGA   1260

Qy   1261 AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGGTGGATGGCCGGCAAGGT   1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGGTGGATGGCCGGCAAGGT   1320

Qy   1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCACAAGTTGCGACTAT   1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCACAAGTTGCGACTAT   1380

Qy   1381 GTTTAGCAGTGTACACTACCTGGCGGTTGGCGCTCTGATCTCGATCTACTACATGGGCAA   1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GTTTAGCAGTGTACACTACCTGGCGGTTGGCGCTCTGATCTCGATCTACTACATGGGCAA   1440

Qy   1441 GTGGTATCAGTTGCTCCTAGCGGTTTATGCTTTACATAGAGCGACCCTCTGGAAACCCCAT   1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GTGGTATCAGTTGCTCCTAGCGGTTTATGCTTTACATAGAGCGACCCTCTGGAAACCCTAT   1500

Qy   1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCCTTGATGATACCATGTCC   1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCCTTGATGATACATGTCC   1560

Qy   1561 TTGCCACTCTTATTTGAGTGAAGATGTGTCAGAGAGTCATTGGTTGTACAGTCCAAAGTGGAC   1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TTGCCACTCTTATTTGAGTGAAGATGTGTCAGAGAGTCATTGGTTGTACAGTCCAAAGTGGAC   1620

Qy   1621 CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTG   1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CAGGCCTATCACTCGTCACTCTAGAGTATATAACAACTCCATATCTTGGTACCCCTATACAATCCCTG   1680

Qy   1681 TGCGAGGGGATGGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA   1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGCGAGGGGATGGTATGGTTAAATTCAAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA   1740

Qy   1741 TGTGCCATCGTACTGCACTATGGGCACTGATGCAGTGTGGAACGACACTGCAACACTTA   1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TGTGCCATCGTACTGCACTATGGGCCACTGATGCAGTGTGGAACGACACTGCAACACTTA   1800

Qy   1801 CGAAGCATGCGGTGTAACACCGCATGGCCACATGGCCGCACACCGGCTCAGCCCTGAA   1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db  1801  CGAAGCATGCGGTGTAACACCACCGGATCGGCTAACAACCGCGATCGGCACAACGGCTCAGCCCTGAA  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1861  ATTGGCTATATTACAATACCCCTGGGTTTACTTTATGACCCACCTTGGATGTC                 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  ATTGGCTATATTACAATACCCCTGGGTTTACTTTATGACCCACCTTGGATGTC                 1920

Qy  1921  AGGCCATTTGTATTTGATTTGGAGGGATCAGATACCCCTATAGTTTATGACCCTGTGAA           1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  AGGCCATTTGTATTTGATTTGGAGGGATCAGATACCCCTATAGTTTATGACCCTGTGAA           1980

Qy  1981  TTCCACTCTCCTACCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCTGTGGTACG          2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  TTCCACTCTCCTACCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCTGTGGTACG          2040

Qy  2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTCGCAGTGATGTGAAAGACCTAGCACCAGAG         2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTCGCAGTGATGTGAAAGACCTAGCACCAGG          2100

Qy  2101  ATTGATCACCAAAGACAAAGCCTGGAAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC         2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  ATTGATCACCAAAGACAAAGCCTGGAAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC         2160

Qy  2161  TTTGTCTCTTACGGGGAGTTACCACCAAGGCCCGTGGTCTAATTCTCGTGTGGGGTGTGTGG        2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TTTGTCTCTTACGGGGAGTTACCACCAAGGCCCGTGGTCTAATTCTCGTGTGGGGTGTGTGG        2220

Qy  2221  CAGCAAGTATCTTATTTAGCCTACCTCGTGTACTTGTTTGGGCCGCGCTTC                   2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  CAGCAAGTATCTTATTTAGCCTACCTCGTGTACTTGTTTGGGCCGCGCTTC                   2280

Qy  2281  TGGTTACCCTTTGCGTCTGTGTATCCCAGTCGTATCTCCAAGCTGGCTGGGATGT              2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TGGTTACCCTTTGCGTCCGTGTATCCCAGTCGTATCTCCAAGCTGGCTGGGATGT              2340

Qy  2341  TTTGTCTAAAGCTCAAGTAGCTTCCTTTGCTTTGATTTTCTTCATCTGTTGTATCTCCG           2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  TTTGTCTAAAGCTCAAGTAGCTTCCTTTGCTTTGATTTTCTTCATCTGTTGTATCTCCG           2400

Qy  2401  CTGCAGGCTACGTTATGCGTCCTCGAGCTGGCCCATGGCGTATCCCAGTCGTTGTGCCCCT         2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  CTGCAGGCTACGTTATGCGTCCTCGAGCTGGCCCATGGCTTGTGCCCCTTGTGCCCCT           2460

Qy  2461  AACTTTCTTGTTGCAGCAGCTGCTGCGCCAACCAGATTATGACTGGTGGGTGCGACTGCT          2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  AACTTTCTTGTTGCAGCAGCTGCTGCGCCAACCAGATTATGACTGGTGGGTGCGACTGCT          2520

Qy  2521  AGTGGGCAGGGTTAGTTTTGTGGGCCGCGGCGTAACCGGTCACACGGCATAGCTCTGCTTGT       2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  AGTGGGCAGGGTTAGTTTTGTGGGCCGCGGCGTAACCGGTCACACGGCATAGCTCTGCTTGT       2580

Qy  2581  AGGTCCTTGGCCCTCGGATAGTTGGTTTAACCCTTGGTTACGGCCTTGGTTACGCTTCAGC        2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  AGGTCCTTGGCCCTCGGATAGTTGGTTTAACCCTTGGTTACGCCCTGCTTCAGC             2640

Qy  2641  TTTTGATACCGAGATAATTGGAGGGCTGACATACCACCTGTAGTAGCATTAGTTGTCAT          2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  TTTTGACACCGAGATAATTGGAGGGCTGACATACCACCTGTAGTAGCATTAGTTGTCAT          2700

Qy  2701  GTCTCGTTTGGCTTCTTGCTTCACTTGTTACCTCGCTGGTGCTTTCGTCGTGTGTCT           2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  GTCTCGTTTGGCTTCTTGCTTCACTTGTTACCTCGCTGGTGCTTTCGTCGTGTGTCA           2760

Qy  2761  TTGGCAACGTTGGGAGAATTGGTTTGGAACGTTACACTAAGACCGGAGAGGTTTTCCT           2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  TTGGCAACGTTGGGAGAATTGGTTTGGAACGTTACACTAAGACCGGAGAGGTTTCCT           2820

Qy  2821  TGTGCTCGGTTTGTTCCCCGGTGCGACATATGCGCGGTGACTTTCTGTGTGTGTGTCA          2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  TGTGCTCGGTTTGTTCCCCGGTGCGACATATGCGCGGTGACTTTCTGTGTGTGTGTCA          2880

Qy  2881  CGTAGCTCTTCTATGTTTAAACATGTGCAGCATCCAGTGCAGTCGTTCTTTGGGACTCTAGGGT    2940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  CGTAGCTCTTCTATGTTTAACATCCAGTGCAGCATCCAGTGCAGTCGTTCTTTGGGACTCTAGGGT  2940
```

```
Qy  2941  TAGGGCCCATAGAAGTTGGTGCGTCTCGAAAGTGTCATGCTTGGTATTCTCATTATGT            3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941  TAGGGCCCATAGAAGTTGGTGCGTCTCGAAAGTGTCATGCTTGGTATTCTCATTATGT            3000

Qy  3001  TCTTAAGTTTTCCTCTCTTAGTGTTTGGTGAGAATGGTGTTTTCTATAAGCACTTGCA            3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  TCTTAAGTTTTCCTCTCTTAGTGTTTGGTGAGAATGGTGTTTTCTATAAGCACTTGCA            3060

Qy  3061  TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCAAGAGCCATTTTCCC            3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061  TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCAAGAGCCATTTTCCC            3120

Qy  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACGCTTGGGCGTGTGGGGACAC         3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACGCTTGGGCGTGTGGGGACAC         3180

Qy  3181  GGTTGATGGTTTGCCCGTTGTTGCCCGTGTTGCCCGTGTTTTGTTTTCGGCAGGGGTTAGCTAT      3240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181  GGTTGATGGTTTGCCCGTTGTTGCCCGTGTTGCCCGTGTTTTGTTTTCGGCAGGGGTTAGCTAT      3240

Qy  3241  GCCGCCAGATGGGTGGGCCATTACCGACACCTTTACGCTGCAGTGTCTCTCTGGAACGTGG         3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241  GCCGCCAGATGGGTGGGCCATTACCGACACCTTTACGCTGCAGTGTCTCTCTGAACGTGG          3300

Qy  3301  CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC           3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301  CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC           3360

Qy  3361  TATCTTCAGATTAGGATCTCTCGGCCACTAGCTACACATGGATTTGTTGTGTGACAACGTGTT      3420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361  TATCTTCAGATTAGGATCTCTCGGCCACTAGCTACACATGGATTTGTTGTGTGACAACGTGTT      3420

Qy  3421  GTATACTGCTCACCATGGCAGCAAGGGGCGCCGGTTGGCTCATCCCACCAGGCTCTATACA        3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421  GTATACTGCTCACCATGGCAGCAAGGGGCGCCGGTTGGCTCATCCCACCAGGCTCTATACA        3480

Qy  3481  CCCAATAACCGTTGACGCGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG          3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481  CCCAATAACCGTTGACGCGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG          3540

Qy  3541  GTCCCTTACTCGGTGCGTCTTGCGCGGGGAGACCAAGGGTATCTGTAACACGACTGGGGTC        3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541  GTCCCTTACTCGGTGCGTCTTGCGCGGGGAGACCAAGGGTATCTGTAACACGACTGGGGTC        3600

Qy  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCCTTATTGGTGTGTGCGGGGCCCTTCCCAT          3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCCTTATTGGTGTGTGCGGGGCCCTTCCCAT          3660

Qy  3661  GGCTGTTGCCAAGGGTTCTTCAGGTGCCCGATTCTGTTGCCCGATTCTGTTGCCCGATTGTTATTGG  3720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661  GGCTGTTGCCAAGGGTTCTTCAGGTGCCCGATTCTGTTGCCCGATTCTGTTGCCCGATTATTGG     3720

Qy  3721  GATGTTCACCGCTGCTAGAAATTCTGGCGGTTCAGTAGCATTAGGGTTAGGCCGTT             3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721  GATGTTCACCGCTGCTAGAAATTCTGGCGGTTCAGTCAGCCAGATTAGGGTTAGGCCGTT          3780

Qy  3781  GGTGTGTGCTGATACCATCCCCAGTACAGCACATGCCACTCTTGATACAAAACCTAC             3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781  GGTGTGTGCTGATACCATCCCCAGTACAGCACATGCCACTCTTGATACAAAACCTAC             3840

Qy  3841  TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAAGTGCAAGTCAAC           3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841  TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGCAAGTCAAC           3900

Qy  3901  CAAATTACCACTTTCTTACACGTGCAGGAGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT        3960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901  CAAATTACCACTTTCTTACACGTGCAGGAGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT        3960

Qy  3961  GGCTACACACAGCATCAATTGCCAAAGTACAATGCCGACACGCGACTACGCCGTGAATCCAAATTG     4020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961  GGCTACACACAGCATCAATTGCCAAAGTACAATGCCGACACGCGACTACGCCGTGAATCCAAATTG     4020
```

```
Qy   4021  CTATTTTAATGGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGACACATATGGCAT  4080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021  CTATTTTAATGGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGACACATATGGCAT  4080

Qy   4081  GTACCTGACCGGAGCCATGTTCCCGGAACTATGATGGGCATTGGCAAGGTCTCATCCAGC   4140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081  GTACCTGACCGGAGCCATGTTCCCGGAACTATGATGGGCATTGGCAAGGTCTCATCCAGC   4140

Qy   4141  TACCGATGCAACCACCCGTGTTGGGCATTGGAAAAGGTCCTAACCGAAGCTCATCCAAAAA  4200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141  TACCGATGCAACCACCCGTGTTGGGCATTGGAAAAGGTCCTAACCGAAGCTCATCCAAAAA  4200

Qy   4201  TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCCTGGAGTAATCCCTACCACCACATGC  4260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201  TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCCTGGAGTAATCCCTACCACCACATGC  4260

Qy   4261  CAACTAACTACTGAGATTCAATTAACCGATGAGGGCACTATCCCCTTTCATGGAAAAAAGAT  4320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261  CAACTAACTACTGAGATTCAATTAACCGATGAGGGCACTATCCCCTTTCATGGAAAAAAGAT  4320

Qy   4321  TAAGGAGAAAATCTGAAGAAAGGGAGACACTTATCTTTGAGGCTACCAAAAAACACTG   4380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321  TAAGGAGAAAATCTGAAGAAAGGGAGACACTTATCTTTGAGGCTACCAAAAAACACTG   4380

Qy   4381  TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTTACTATAGGGG  4440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381  TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTTACTATAGGGG  4440

Qy   4441  ATGTGACATCTCAAAATCCCTGAGGGCGACTGTGTAGTAGTCCACTGATGCCTTGTG   4500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441  ATGTGACATCTCAAAATCCCTGAGGGCGACTGTGTAGTAGTCCACTGATGCCTTGTG   4500

Qy   4501  TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACGTGCAGCCTCATGGTAGAGGCAC   4560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501  TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACGTGCAGCCTCATGGTAGAGGCAC   4560

Qy   4561  ATGCCATGGTGACCCTGACCCTACTTTCACCATGGGTGTTCGTGTGTGCGGGTTTCAGC   4620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561  ATGCCATGGTGACCCTGACCCTACTTTCACCATGGGTGTTCGTGTGTGCGGGTTTCAGC   4620

Qy   4621  AATAGTTAAAGGCCAGCAGCCGCACAGCCCGTGGGAGAGAGTCGGCATATACTACTA   4680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4621  AATAGTTAAAGGCCAGCAGCCGCACAGCCCGTGGGAGAGAGTCGGCATATACTACTA   4680

Qy   4681  TGTAGACGGGAGTTGTACCCCTTCGGGTATGTTCCTGAAGTCAACATTGTTGAAGCCTT   4740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4681  TGTAGACGGGAGTTGTACCCCTTCGGGTATGTTCCTGAAGTCAACATTGTTGAAGCCTT   4740

Qy   4741  CGACCGCAGCCAAGGCATGGTATGGTTTGTCATCAACAGAGCTCAAACTATTCTGGACAC   4800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741  CGACCGCAGCCAAGGCATGGTATGGTTTGTCATCAACAGAGCTCAAACTATTCTGGACAC   4800

Qy   4801  CTATCGCACCCACCCAACCTGGGTTACCTGGGTACTGGAATAGGAGCAAATTTGGACGAGTGGCTGATCT  4860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801  CTATCGCACCCACCCAACCTGGGTTACCTGGGTACTGGAATAGGAGCAAATTTGGACGAGTGGCTGATCT  4860

Qy   4861  CTTTTCTATGGTCAACCCGGACCCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA   4920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861  CTTTTCTATGGTCAACCCGGACCCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA   4920

Qy   4921  TTATGTTTTGTGACTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGTGCTCC   4980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921  TTATGTTTTGTGACTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGTGCTCC   4980

Qy   4981  CAATGACGACCACCGTTGGCAGGAGAGCCCGGCTTGGGAAAAAACCTTGTGTGGGTTCTGTG  5040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981  CAATGACGACCACCGTTGGCAGGAGAGCCCGGCTTGGGAAAAAACCTTGTGTGGGTTCTGTG  5040

Qy   5041  GCGCTTGGACGCGCTGCAGCCGCGCTGTCCTGGCCCAGAGCCCAGCCGAGGTGACCAGATACCA  5100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041  GCGCTTGGACGCGCTGCAGCCGCGCTGTCCTGGCCCAGAGCCCAGCCGAGGTGACCAGATACCA  5100

Qy   5101  AATGTGCTTCACTGAAGTCAATATACTTCTGGGACAGCCGCACTCCGCTGTGTTGGCGTTCGGAGT  5160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101  AATGTGCTTCACTGAAGTCAATATACTTCTGGGACAGCAGCCGCACTCCGCTGTGTGGCGGCGTTGCTGCGTC  5220
```

```
Qy   5161  GGCTATGGCTTATCTAGCCAATGAACTTTGGCGGCCACTTGTGTGCGGCGTTGCTGCGTC  5220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161  GGCTATGGCTTATCTAGCCAATGAACTTTGGCGGCCACTTGTGTGCGGCGTTGCTGCGTC  5220

Qy   5221  TATTACATCAGTCCTCACCGGTGCTACCGTGTCCGCCCCCAGTGGTTGACGAAGAAAATCGT  5280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221  TATTACATCAGTCCTCACCGGTGCTACCGTGTCCGCCCCCAGTGGTTGACGAAGAAAATCGT  5280

Qy   5281  GGAGGAGTGTGCATCATTCATTCCCTGGAGGCATGTTGCTGCAATTGACAAGCTGAA  5340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5281  GGAGGAGTGTGCATCATTCATTCCCTGGAGGCATGTTGCTGCAATCGATAAGCTGAA  5340

Qy   5341  GAGTACAATCACCACACTAGTCCTTTCACATTGGAAACCGCCCTTGAAAAACTTAACAC  5400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5341  GAGTACAATAACCACACCACACTAGTCCTTTCACATTGGAAACCGCCCTTGAAAAACTTAACAC  5400

Qy   5401  CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGGTTTAGT  5460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5401  CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGGCTTAGT  5460

Qy   5461  CACTTTACCTGACAATCCCTTGCATCATGCAGTGTTTTGCTTTCATTGCGGGTATTACTAC  5520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5461  CACTTTACCTGACAATCCCTTGCATCATGCAGTGTTTTGCTTTCATTGCGGGTATTACTAC  5520

Qy   5521  CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTTGGAGGCGGCAATTGCGTCCAA  5580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5521  CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTTGGAGGCGGCAATTGCGTCCAA  5580

Qy   5581  GCTTACAGACGCTAGAGGCGCACTGGCGTTCATGATGGCCGGGGCCTGCGGGAACAGCTCT  5640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5581  GCTTACAGACGCTAGAGGCGCACTGGCGTTCATGATGGCCGGGGCCTGCGGGAACAGCTCT  5640

Qy   5641  TGGTACATGGACATCGGTGGGTTTTGTCTTTGACATGGGTCTATGCTCGCCGCCTC  5700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5641  TGGTACATGGACATCGGTGGGTTTTGTCTTTGACATGGGTCTATGCTCGCCGCCTC  5700

Qy   5701  ATCCACTGCTTGCTTGCTTGACATTTAAATGCTTAAATGTTGATGGGTGAGTTCGGCATCAGCT  5760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5701  ATCCACTGCTTGCTTGCTTGACATTTAAATGCTTAAATGTTGATGGGTGAGTTCGGCATCAGCT  5760

Qy   5761  TGCTGGTTAGTCTACTCCGGCGTTCAATCCGGCCCGCAGAGTTGTGTGTCGTCTTGTCAGC  5820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5761  TGCTGGTTAGTCTACTCCGGCGTTCAATCCGGCCCGCAGAGTTGTGTGTCGTCTTGTCAGC  5820

Qy   5821  TTTGTCAATGTTGCTTGCTGACAACAGCAGGCCCAGATCACTCTGCCCACAACAGACTTCTTAC  5880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5821  TTTGTCAATGTTGCTTGCTGACAACAGCAGGCCCAGATCACTCTGCCCACAACAGACTTCTTAC  5880

Qy   5881  TATGCTTGCTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG  5940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5881  TATGCTTGCTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG  5940

Qy   5941  CAGGAAGAGATACTGGCATTCTCGGAGCATTCTCGGAGGCATCATATATCAGCTTGCAT  6000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5941  CAGGAAGAGATACTGGCATTCTCGGAGCATTCTCGGAGGCATCATATATCAGCTTGCAT  6000

Qy   6001  CCGTTGGCTCCACACCCGACGGAGGATGATTGCGGCCTCATTGCTTGTGCTTCTTAGAGAT  6060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6001  CCGTTGGCTCCACACCCGACGGAGGATGATTGCGGCCTCATTGCTTGTGCTTCTTAGAGAT  6060

Qy   6061  TTGGCAGTATGTGCAATTTCTTTGATTTGTGATTTGTGATTTTTAAAGTCCTTAAAGCTGGAGTTCA  6120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6061  TTGGCAGTATGTGCAATTTCTTTGTGCAATTTCTTTGTGATTTGTGATTTTTAAAGCTGGAGTTCA  6120

Qy   6121  GAGCATGGTTAACATTCCTGTCCTTTCAGTCGTCCTTTCTCACAGCTGCCAGAAGGGGGTACAAGGGCCC  6180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6121  GAGCATGGTTAACATTCCTGTCCTTTCGTGTCCTTTCTCACAGCTGCCAGAAGGGGTCAAGGGGCCC  6180

Qy   6181  CTGGATTGCTTCACTGAAGTCAATATGCTCCAAGCAGCGCCACTCGCTGTGTTGGCGTTCGGAGT  6240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
                                                                                              7380
Qy  7321 TCTGTCTGCAACTCGGGCCATCACTACTAGTGGTTTCCTCAAACAACAAGATCATTGGTGTATGT 7380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7321 TCTGTCTGCAACTCGGGCCATCACTACTAGTGGTTTCCTCAAACAACAAGATCATTGGTGTATGT 7380

Qy  7381 GACTGAGCGCGGATGGGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCT 7440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381 GACTGAGCGCGGATGGGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCT 7440

Qy  7441 GTTCCCCATCATCACCAACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441 GTTCCCCATCATCACCAACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500

Qy  7501 CGGTGTCATGTGGGACTATGATGAAGTAGCAGTCACACGCCCTCTAAGTCTGCTAAGTC 7560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7501 CGGTGTCATGTGGGACTATGATGAAGTAGCAGTCACACGCCCTCTAAGTCTGCTAAGTC 7560

Qy  7561 CCACATCACTGGCCTTCGGGGCACTGATGTTCGTTCGTTCTGGAGCAGCCCGCAAGGCTGTTCT 7620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7561 CCACATCACTGGCCTTCGGGGCACTGATGTTCGTTCGTTCTGGAGCAGCCCGCAAGGCTGTTCT 7620

Qy  7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGGAGTCATTATCGGCAAACTGTGAT 7680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGGAGTCATTATCGGCAAACTGTGAT 7680

Qy  7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAAACAACAAAGAAACCCCCAAG 7740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAAACAACAAAGAAACCCCCAAG 7740

Qy  7741 GCTTATCTCGTACCCCACCTTGAAATGAGAGTGTTGAGAAGATGTACTACGGTCAGGT 7800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7741 GCTTATCTCGTACCCCACCTTGAAATGAGAGTGTTGAGAAGATGTACTACGGTCAGGT 7800

Qy  7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGTGGCGTACGGGTTTGTAGATCCACGTAC 7860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGTGGCGTACGGGTTTGTCGACCCACGTAC 7860

Qy  7861 CCGTGTCAAGCGTCGTGTTGTCGATGTGGTTCACCCGGATGCGGAGCCAGTCGGACCATGCGGATAC 7920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7861 CCGTGTCAAGCGTCGTGTTGTCGATGTGGTTCACCCGGATGCGGAGCCAGTCGGACCATGCGGATAC 7920

Qy  7921 AGTGTGTTTGACAGTAGCCATCACACACCGGAGGATATCATGGTGGAGACAGACATCTACTC 7980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7921 AGTGTGTTTGACAGTAGCCATCACACACCGGAGGATATCATGGTGGAGACAGACATCTACTC 7980

Qy  7981 AGCAGCTAAACTCAGTGACCAACACCGGCGATTCACCACCATTGCGAGGCAGTTATA 8040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7981 AGCAGCTAAACTCAGTGACCAACACCGGCGATTCACCACCATTGCGAGGCAGTTATA 8040

Qy  8041 CGCTGGGAGGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGTGTAGGTC 8100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8041 CGCTGGGAGGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGTGTAGGTC 8100

Qy  8101 TTCCGGCGTCTATACTACCTCAAGTTCCAACAGTTGACCTGCTGGCTGCTGAAGGTAAATGC 8160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8101 TTCCGGCGTCTATACTACCTCAAGTTCCAACAGTTGACCTGCTGGCTGCTGAAGGTAAATGC 8160

Qy  8161 TGCAGCCGAACAGGCTGGCATGGAAGAACCCCGCGTTCCTTATTTGCGGCGATGATTGCAC 8220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161 TGCAGCCGAACAGGCTGGCATGGAAGAACCCCGCGTTCCTTATTTGCGGCGATGATTGCAC 8220

Qy  8221 CGTAATTTGGAAGAGCGCCGAGCAATGCGTCTTTGCTAG 8280
         |||||||||||||||||||||||||||||||||||||||
Db  8221 CGTAATTTGGAAGAGCGCCGAGCAATGCGTCTTTGCTAG 8280

Qy  8281 CTGGATGAAGGTGATGGGTGCACCACCAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281 CTGGATGAAGGTGATGGGTGCACCACCAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340

Qy  8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400
```

```
Db  6181 CTGGATTGGATCAGTAGTGATGTCTCCAAGCACGCTGTCTGCCATGCGGTGCTGAACTCATCTTTC 6240

Qy  6241 TGTTGAGAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTCGGAG 6300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241 TGTTGAGAATGGTTTGCAAAACTTTACAAAGGACCCAGAACTTGTTCAAATTACTCGGAG 6300

Qy  6301 AGGGGCTGTTCCAGTCAACGCTAGGCGTGTGGGTCGGCTAGACCGGACCCAACTCGATTG 6360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301 AGGGGCTGTTCCAGTCAACGCTAGGCGTGTGGGTCGGCTAGACCGGACCCAACTCGATTG 6360

Qy  6361 GACTAGTTCGTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGGGAGAAATGGGAGA 6420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361 GACTAGTTCGTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGGAAATGGGAGA 6420

Qy  6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCAAC 6480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCCAAC 6480

Qy  6481 CTTGAGAGCTGCAGTGGCCGTGACGGCGTACAGGTTATCCTAGTGTATCTAGGTGAGCCCAA 6540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481 CTTGAGAGCTGCAGTGGCCGTGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540

Qy  6541 AACTCCTTGGACGACATCTGCTTGCTGTCTGACGGTAAGGGTAAAACTGTTAA 6600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541 AACTCCTTGGACGACATCTGCTTGCTGTCTGACGGTAAGGGTAAAACTGTTAA 6600

Qy  6601 GCTTCCCCTCCGCGGTTGACGACCACCAAACTGCCAGCCCTCCAGCGCAACTTAATTTGCGTGA 6660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601 GCTTCCCCTCCGCGGTTGACGACCACCAAACTGCCAGCCCTCCAGCGCAACTTAATTTGCGTGA 6660

Qy  6661 TGCACTTGAGACAAATGACTGTAATTCCACAAACACTCCTAGTGTTGAAGCCGCAGT 6720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661 TGCACTTGAGACAAATGACTGTAATTCCATAAACAACACTCCTAGTGTTGAAGCCGCAGT 6720

Qy  6721 GTCCGGCTCTTGTTTCAAACAGAAGTTGCGGCGTACAAACAATTGCTTGAGGCAATTTC 6780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721 GTCCGGCTCTTGTTTCAAACAGAAGTTGCGGCGTACAAACAATTGCTTGAGGCAATTTC 6780

Qy  6781 AGGAGTGGCGGTTGACCACCACCAAACTGCCGCAGCGCCCTCCATCATGCGAAGAGGTAGTGTAAGAAA 6840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781 AGCTGGGCGGTTGACCACCACCAAACTGCCGCAGCGCCCTCCATCATGCGAAGAGGTAGTGTAAGAAA 6840

Qy  6841 GCGCCAGTTCCGGGCAAGAACAGGCCGAAAACAGGCGGCGGCGGAGGCGGAGGCCGTCCC 6900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841 GCGCCAGTTCCGGGCAAGAACAGGCCGAAAACAGGCGGCGGCGGAGGCGGAGGCCGTCCC 6900

Qy  6901 AGGAGTGTCATGTCGAAAGCTGAAACGGTCAGCGAAGGTCCTTCGAGGCAATTTC 6960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901 AGGAGTGTCATGTCCGAAAGCTGAAACGGTCAGCGAAGGTCCTTCGAGGCAATTTC 6960

Qy  6961 CCCTCCTTCACAGTTCTTCCACCTGGCCCCATGCCCTGCGTTGGGAGCGGCGGTGA 7020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961 CCCTTCTTCACCACCCTGTTCTTCCACCTGGCCCCATGCCCTGCGTTGGGAGCGGCGGTGA 7020

Qy  7021 GTGTAACCCTTTCACTGCGAATTGGATGTGTCAATGACCGGAAAACAGGCGGCCCTGATGA 7080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021 GTGTAACCCTTTCACTGCGAATTGGATGTGTCAATGACCGGAAAACAGGCGGCCCTGATGA 7080

Qy  7081 TTTACCCAGTTACCCTCCCAAAAAGGGTCTCTGAATGGTCAGACGAAGTTGGTCGAC 7140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081 TTTACCCAGTTACCCTCCCAAAAAGGGTCTCTGAATGGTCAGACGAAGTTGGTCAAC 7140

Qy  7141 GGCTACAACCGCTTCCAGGTTCTACAGTCTACTGGCTACACCGTAACGCGGGGAAAAGGA 7200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141 GACTACAACCGCTTCCAGTCCTACACGGTCTACTTAAGATACGGGGAAAAGGA 7200

Qy  7201 TTCCACTTCAGCCAGCTTCACTGCAATGAGAGTTGGAAAGAGTGAGTT 7260
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201 TTCCACTTCAGCCAGCTTCACTGCAATGAGAGTTGGAAAGAGTGAGTT 7260

Qy  7261 TTCGTGCAGCATGAGCTACAACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261 TTCGTGCAGCATGAGCTACAACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320
```

```
FT                        /tag= b
FT                        /codon_start= 446. .448
XX
PN   WO9707224-A1.
XX
PD   27-FEB-1997.
XX
PF   14-AUG-1996;    96WO-US013198.
XX
PR   14-AUG-1995;    95US-0002265P.
PR   21-DEC-1995;    95US-00580038.
PR   19-APR-1996;    96US-00639857.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Simons JN, Desai SM, Mushahwar IK;
XX
DR   WPI; 1997-165306/15.
XX
PT   Controlling translation of hepatitis GB viral nucleic acid - with non-
PT   natural hybridising sequence, used for treatment of infection, screening
PT   blood etc.
XX
PS   Disclosure; Page 61-66; 86pp; English.
XX
CC   5' Nontranslated regions (NTRs) (AAT59784-86) were identified for
CC   hepatitis GB virus (HGBV) types A, B and C. Unlike HGBV-A and -C, HGBV-B
CC   appears similar to the hepatitis C virus (HCV) and pestivirus genera of
CC   the Flaviviridae. Similar to HCV, HGBV-B 5'NTR contains an internal
CC   ribosome entry site. Nucleic acids (including antisense) derived from
CC   HGBV 5'NTRs can be used to control the translation of HGBV nucleic acids
CC   to proteins. Blocking or decreasing translation may decrease the
CC   pathology of a viral infection. Enhancement of translation may allow for
CC   stronger immune responses. HGBV nucleic acids can also be used to screen
CC   blood/organs for the presence of HGBV, in epidemiological studies and
CC   possibly to purify HGBV proteins for use in diagnostic assays. (Updated
CC   on 27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 9143 BP; 2071 A; 2266 C; 2349 G; 2457 T; 0 U; 0 Other;
```

```
Query Match             96.2%;  Score 9037.4;  DB 2;  Length 9143;
Best Local Similarity   99.4%;  Pred. No. 0;
Matches 9082;  Conservative   0;  Mismatches  51;  Indels   4;  Gaps   1;

Qy         1 ACCACAAACACTCCAGTTGTTACACTCCGTCTAGGAATGCTCTCCGGAGCACCCCCTAG 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ACCACAAACACTCCAGTTGTTACACTCCGTCTAGGAATGCTCTCCGGAGCACCCCCTAG 60

Qy        61 CAGGGGCGTGGGGATTCCCCTGCCCCGCTCTGCAGAAGGTGGAGCCAACCACCTTAGTAT 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 CAGGGGCGTGGGGATTCCCCTGCCCCGCTCTGCAGAAGGTGGAGCCAACCACCTTAGTAT 120

Qy       121 GTAGGGCGCGGGACTCCGGTGGCTCGCGGTCGCGATGACAAGCGCCAAGCTTGACTTGGATGGC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GTAGGGCGCGGGACTCCGGTGGCTCGCGGTCGCGATGACAAGCGCCAAGCTTGACTTGGATGGC 180

Qy       181 CCTGATGGGCGTTCATGGGTTCGGTGGCTTTAGGCAGCCTTAGGCAGCGCCCACGCCCCACCA 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CCTGATGGGCGTTCATGGGTTCGGTGGCTTTAGGCAGCGCCTTAGGCAGCGCCCACGCCCCACCA 240

Qy       241 CCTCCCAGATAGAGCGGCGGCGGGCCACTGTAGGGAAGACCGGGGAAGACGGGGACCGGCGTCACTACCAAGGACG 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CCTCCCAGATAGAGCGGCGGCGGGCCACTGTAGGGAAGACCGGGGAAGACGGGGACCGGCGTCACTACCAAGGACG 300

Qy       301 GTAGGGCCGGCGGGACTCATGACGCTCGCGGTCGCGATGACAAGCGCCAAGCTTGACTTGGCCACCTATATGTGT 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 GTAGGGCCGGCGGGACTCATGACGCTCGCGGTCGCGATGACAAGCGCCAAGCTTGACTTGGCCACCTATATGTGT 360

Qy       361 TGGGATGGTTGGGGTTCCTGATATCACGCTACGATATCACGCTGCCTATAGGGTCCTTGCGAGGGGAT 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 TGGGATGGTTGGGGTTCCTGATATCACGCTGCCTATCACGCTGCCTGATAGGGTCCTTGCGAGGGGAT 420

Qy       421 CTGGGAGTCTCTAGACCGTAGCACATGCCGTTGTATTTCTACTCAAACAAGTCCTGTACC 480
```

```
Qy      8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCTTGGCAGTGCTCTGCCGAGGGTCTGGG 8460
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCTTGGCAGTGCTCTGCCGAGGGTCTGGG 8460

Qy      8461 ATACAACCCAGTGCGTCGGTGGATTGGGTATCTAATACATCACTACCCATGTTTGTGGGT 8520
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8461 ATACAACCCAGTGCGTCGGTGGATTGGGTATCTAATACATCACTACCCATGTTTGTGGGT 8520

Qy      8521 TAGCCGTGTGTGGCGTGTCCATTTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8521 TAGCCGTGTGTGGCGTGTCCATTTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580

Qy      8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640

Qy      8641 CATCATTGCTGGTGGTGCACGGTATTGAGGCTTTCTCGGTGGTGCGTACCAACGCTGA 8700
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8641 CATCATTGCTGGTGGTGCACGGTATTGAGGCTTTCTCGGTGGTGCGTACCAACGCTGA 8700

Qy      8701 GATCCTCAGAGTTCCAATCACTAACAGACATGACCATGCCCCCCTCGCAGCCTGGCG 8760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8701 GATCCTCAGAGTTCCAATCACTAACAGACATGACCATGCCCCCCTCGCAGCCTGGCG 8760

Qy      8761 AAAGAAAGCCAGGGGCGGTCCTCGCCAGCGCCAAGAGGCGTGGCGAGCACACGCAAAATT 8820
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8761 AAAGAAAGCCAGGGGCGGTCCTCGCCAGCGCCAAGAGGCGTGGCGAGCACACGCAAAATT 8820

Qy      8821 GGCTTCGCTTCCTTCCTCTGGCATGCTACATCTAGACCTTCTACCAGATTGGATAAGACGAG 8880
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8821 GGCTTCGCTTCCTTCCTCTGGCATGCTACATCTAGACCTTCTACCAGATTGGATAAGACGAG 8880

Qy      8881 CGTGGCTCGGTACACCACTTTCAATTATGTGATGTTTACTCCCCGGAGGGGATGTGTT 8940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8881 CGTGGCTCGGTACACCACTTTCAATTATGTGATGTTTACTCCCCGGAGGGGATGTGTT 8940

Qy      8941 TATTACCACCAGAGAAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTTT 9000
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8941 TGTTTACCACCAGAGAAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTTT 9000

Qy      9001 TGCCCTAGGGCTCATTGCTGTTGGACTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA 9060
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9001 TGCCCTAGGGCTCATTGCTGTTGGACTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA 9060

Qy      9061 CTAACAG----TTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC 9116
             |||||||     ||||||||||||||||||||||||||||||||||||||||||||
Db      9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC 9120

Qy      9117 GGGCTTAACGACCCCGC 9133
             |||||||||||||||||
Db      9121 GGGCTTAACGACCCCGC 9137
```

```
RESULT 9
AAT59785
ID   AAT59785 standard; DNA; 9143 BP.
XX
AC   AAT59785;
XX
DT   27-AUG-2003 (revised)
DT   18-MAY-1997 (first entry)
XX
DE   Hepatitis GB virus-B 5'nontranslated region.
XX
KW   HGBV; flavivirus; translation; antisense; therapy; blood screening;
KW   diagnosis; SB.
XX
OS   Hepatitis GB virus B.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..445
FT                   /tag= a
FT   CDS             446. .448
```

Left column (positions 421–1560):

```
Db  421  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  480
         CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACC

Qy  481  TGCGCCCAGAACCGCCAAGAACAAGCAGCAGCAGGCTTCTATATCCTGTTCCATTAAAAC  540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  TGCGCCCAGAACCGCCAAGAACAAGCAGCAGCAGGCTTCTATATCCTGTTCCATTAAAAC  540

Qy  541  ATCTGTTGAAAGGGGACAACGACAACGAGCAAAGCGCAAAGCGCAAAGTCTCGGCCTCGTAA  600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  ATCTGTTGAAAGGGGACAACGACAACGAGCAAAGCGCAAAGCGCAAAGTCTCGGCCTCGTAA  600

Qy  601  TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC  660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC  660

Qy  661  TCATGGTTGGGGACGCCAAGACCCTCGGCAATCTTGGAATCTTCTGA  720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  TCATGGTTGGGGACGCCAAGACCCTCGGCAATCTTGGAATCTTCTGA  720

Qy  721  TTACCCTTTGGGGTGGATTGGTGATGTTACAACTCACACACTTCTAGTAGGCCCGCTGGT  780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  TTACCCTTTGGGGTGGATTGGTGATGTTACAACTCACACACTTCTAGTAGGCCCGCTGGT  780

Qy  781  GGCAGGAGCGGTCGTTGCGACCAGTCTGCTGCCAGAGTAGTACGGTTGCTGGAGGAGTCAA  840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  GGCAGGAGCGGTCGTTGCGACCAGTCTGCTGCCAGAGTAGTACGGTTGCTGGAGGAGTCAA  840

Qy  841  CTGGGCTACTGGTTGGTTCGGTGTCCCACCTTTTGTGTGGTATGTCTGCTATCTTTGGCCTG  900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  CTGGGCTACTGGTTGGTTCGGTGTCCCACCTTTTGTGTGGTATGTCTGCTATCTTTGGCCTG  900

Qy  901  TCCCTGTAGTGGGCGCGGGGTCACTGACCAGACACAAATACCACAATTCCTGACCAATTG  960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  TCCCTGTAGTGGGCGCGGGGTCACTGACCAGACACAAATACCACAATTCCTGACCAATTG  960

Qy  961  CTGCCAGCGTAATCAGTTATCCATTGGTTCTCCTTCCACTTGCCTACACGAGGCCTGGTTG  1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  CTGCCAGCGTAATCAGTTATCCATTGGTTCTCCTTCCACTTGCCTACACGAGGCCTGGTTG  1020

Qy 1021  TGTGATCTGTGCGACGAGTGCTGGGTTCCCGCCAATCCGTACATCTCACACCTTCCAA  1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  TGTGATCTGTGCGACGAGTGCTGGGTTCCCGCCAATCCGTACATCTCACACCTTCCAA  1080

Qy 1081  TTTGGACTGGCACGGACTCCTTCTTCTGGCTGACCACATTGATTTTGTTATGGGGCTCTTGT  1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081  TTTGGACTGGCACGGACTCCTTCTTCTGGCTGACCACATTGATTTTGTTATGGGGCTCTTGT  1140

Qy 1141  GACCTGTGACGCCCTTGAGCATTGGTGAGTGTGTCGTGTATTAGTCGGTGTACTTGTTACCTGGTTG  1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141  GACCTGTGACGCCCTTGAGCATTGGTGAGTGTGTCGTGTATTAGTCGGTGTACTTGTTACCTGGTTG  1200

Qy 1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGTACTTGTTACCTGA  1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGTACTTGTTACCTGA  1260

Qy 1261  AGTGCCCACTGGAATAGATCCTGGGTTTATCGGGTGGATGGCCGGCAAGGT  1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261  AGTGCCCACTGGAATAGATCCTGGGTTTATCGGGTGGATGGCCGGCAAGGT  1320

Qy 1321  CGAGGCTACGTTATGCGTCAATTGCATGGCTTCACAAGTACCATTACCGCTTATTGCGACTAT  1380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321  CGAGGCTACGTTATGCGTCAATTGCATGGCTTCACAAGTACCATTACCGCTTATTGCGACTAT  1380

Qy 1381  GTTTAGCAGTGTACACTACCTGGCGCGGTTGGCGCTCGATCTACTATGCCTCTCGGGGCAA  1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381  GTTTAGCAGTGTACACTACCTGGCGCGGTTGGCGCTCGATCTACTATGCCTCTCGGGGCAA  1440

Qy 1441  GTGGTATCAGTTGCTCCTAGGCTTATGCTTTACATAGAAGCGCACCTTCTGGAAACCCCAT  1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441  GTGGTATCAGTTGCTCCTAGGCTTATGCTTTACATAGAAGCGCACCTTCTGGAAACCCCAT  1500

Qy 1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCCTTGATGATACCATGTCC  1560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Right column (positions 1501–2640):

```
Db 1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCCTTGATGATACCATGTCC  1560

Qy 1561  TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAGAGTCATTTGTTACAGTGAC  1620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561  TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAGAGTCATTTGTTACAGTGGAC  1620

Qy 1621  CAGGCCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG  1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621  CAGGCCCTGTCACTCGTACTGTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTG  1680

Qy 1681  TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGTTGCTGCCGTATTCGCAA  1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681  TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGTTGCTGCCGTATTCGCAA  1740

Qy 1741  TGTGCCATCGTACTGCACTATGGCACTGATGCGCAGTGTGGAACGACACTGCAACACTTA  1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741  TGTGCCATCGTACTGCACTATGGCACTGATGCGCAGTGTGGAACGACACTGCAACACTTA  1800

Qy 1801  CGAAGCATGCGGTGTAACACCATGCTAACACCGCATGCACAACGGCTCAGCCCTGAA  1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801  CGAAGCATGCGGTGTAACACCATGCTAACACCGCATGCACAACGGCTCAGCCCTGAA  1860

Qy 1861  ATTGGCTATATTACAAATACCCTGGTCTAAAGAAAATGTTAAACCTCATAATTGGATGTC  1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861  ATTGGCTATATTACAAATACCCTGGGTCTAAAGAAAATGTTAAACCTCATAATTGGATGTC  1920

Qy 1921  AGGCCATTGTATTTTGAGGGATCAGATACCCCTATAGTTACTTTATGACCCTGTGAA  1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921  AGGCCATTGTATTTTGAGGGATCAGATACCCCTATAGTTACTTTATGACCCTGTACG  1980

Qy 1981  TTCCACTCTCCTACCACCCGGACCACCAAGGCCGTGGTCAATTCTGTTCCCTTGTGTGG  2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981  TTCCACTCTCCTACCACCCGGACCACCAAGGCCGTGGTCAATTCTGTTCGTGGGTACG  2040

Qy 2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAGG  2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041  TGGTTCTTGGTTACAGGTTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAGG  2100

Qy 2101  ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC  2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101  ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC  2160

Qy 2161  TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGTGCTAATTCTGTTTGGGGTTGTGTGG  2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161  TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGTGCTAATTCTGTTTGGGGTTGTGTGTG  2220

Qy 2221  CAGCAAGTATCTTATTTTAGCCTACTCCTGTACTTGTCCCTTGTTGTTTGGGCCGCCTTC  2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221  CAGCAAGTATCTTATTTTAGCCTACTCCTGTACTTGTCCCTTGTTTGTTTGGGCGCCTTC  2280

Qy 2281  TGGGTTACCCTTGCGTCCTGTGCTCCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGT  2340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2281  TGGGTTACCCTTGCGTCCTGTGCTCCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGT  2340

Qy 2341  TTTGTCTAAAGCTCAAGTAGCTCCCTTTTGCTTTGATTTTCTTCATCGTTGCTATCTCCG  2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2341  TTTGTCTCTAAAGCTCAAGTAGCTCCCTTTTGCTTTGATTTTCTTCATCGTTGCTATCTCCG  2400

Qy 2401  CTGCAGGCTACGTTATGCGTCCCCATGCGTGTCCCATGGCCTTGCCCCT  2460
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401  CTGCAGGCTACGTTATGCGTCCCCATGCGTGTCCCATGGCCTTGCCCCT  2460

Qy 2461  AACTTTCTTTCGTTGCAGCAGCAGATTATGACTAGGCGTGTGGGTGCGACTGCT  2520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2461  AACTTTCTTTCGTTGCAGCAGCAGATTATGACTAGGCGTGTGGGTGCGACTGCT  2520

Qy 2521  AGTGGCAGGGTAGTTTGTTGTGGGCCGGGTCACCGTAACCGTGTCACCCATAGCTCTGCTTGT  2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2521  AGTGGCAGGGTAGTTTGTTGTGGGCCGGGTCGACCGTGTCCAGTCAGTAGCTCTGCTTGT  2580

Qy 2581  AGTTCCTTGGCCTCTGGGTAGCGCTTTAACCCTCTTGGCATTGGCATTTGGCTACGCCTGCTTCAGC  2640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2581  AGTTCCTTGGCCTCTGGGTAGCGCTTTAACCCTCTTGCATTGGCATTTGGCTACGCCTGCTTCAGC  2640
```

```
Qy 2641 TTTTGATACCGAGATAATTGGAGGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT 2700
            |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2641 TTTTGACACCGAGATAATTGGAGGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT 2700

Qy 2701 GTCTCGTTTGGCTTCTTGCTCACTTGTTCCTCGCTGTGTTGTTAGTTAACTCCTATCT 2760
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2701 GTCTCGTTTGGCTTCTTGCTCACTTGTTCCTCGCTGTGTTGTTAGTTAACTCCTATCT 2760

Qy 2761 TTGGCAACGTTGGGAGAATTGGTTTTGGAACGTTACACTAAGACCGGAGACGTTTTCCT 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761 TTGGCAACGTTGGGAGAATTGGTTTTGGAACGTTACACTAAGACCGGAGACGTTTCTCCT 2820

Qy 2821 TGTGCTGGTTGTTTCCCCGGTGCGACATATGACGGCCTGGTGACTTTCTGTGTGTGTCA 2880
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2821 TGTGCTGGTTGTTTCCCCGGTGCGACATATGACGGCCTGGTGACTTTCTGTGTGTGTCA 2880

Qy 2881 CGTAGCTCTTCTATGTTTAACATCCAGTGCAGCATCGTTCTTTGGGACTGACTCTAGGGT 2940
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2881 CGTAGCTCTTCTATGTTTAACATCCAGTGCAGCATCGTTCTTTGGGACTGACTCTAGGGT 2940

Qy 2941 TAGGGCCCATAGAAATGTTGGTGCGTTCGGGAAAGTGTCATGCTTGGTATTCATTCATTATGT 3000
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2941 TAGGGCCCATAGAAATGTTGGTGCGTTCGGGAAAGTGTCATGCTTGGTATTCATTCATTATGT 3000

Qy 3001 TCTTAAGTTTTCCCTCTTAGTTGTTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA 3060
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3001 TCTTAAGTTTTCCCTCTTAGTTGTTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA 3060

Qy 3061 TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCAAGAGCCATTTTTTCCC 3120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3061 TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCAAGAGCCATTTTTTTCCC 3120

Qy 3121 TTTTGAAGGCAAGGCAGGGGTCTATAGGAATGAAGAAGACGCTTGGCGTGTGGGGACAC 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3121 TTTTGAAGGCAAGGCAGGGGTCTATAGGAATGAAGAAGACGCTTGGCGTGTGGGGGACAC 3180

Qy 3181 GGTTGATGGTTTGCCCGTTGTTTCGCGCTCTCGGCGGACCTTGTTTTCGCAGGGTTGCTAT 3240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3181 GGTTGATGGTTTGCCCGTTGTTTCGCGCTCTCGGCGGACCTTGTTTTCGCAGGGTTAGCTAT 3240

Qy 3241 GCCGCCAGATGGTGGGCCATTACCCGCACCTTTACGCAGTGTCTCTCGAACGTGG 3300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3241 GCCGCCAGATGGTGGGCCATTACCCGCACCTTTACGCAGTGTCTCTCGAACGTGG 3300

Qy 3301 CACGCTGCTCAGCGATGGCAGTGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC 3360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3301 CACGCTGCTCAGCGATGGCAGTGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC 3360

Qy 3361 TATCTTCAGATTAGGATCTCTGCCACTAGCTACATGGGGATTTGTTTGTGTCAACGTGTT 3420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3361 TATCTTCAGATTAGGATCTCTGCCACTAGCTACATGGGGATTTGTTTGTGTCAACGTGTT 3420

Qy 3421 GTATACTGCTCACCATGGCAGCAAGGGCGCCGGTTGGCTCATCCCACCAGGCTCCATACA 3480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3421 GTATACTGCTCACCATGGCAGCAAGGGCGCCGGTTGGCTCATCCCACCAGGCTCCATACA 3480

Qy 3481 CCCAATAACCGTTGACCGGTGCTCTTGCGGGAGACCAAGGGATTGTTGTTGTGACAACGTGTT 3540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3481 CCCAATAACCGTTGACCGGTGCTCTTGCGGGAGACCAAGGGATTGTTGTTGTGACAACGTGTC 3540

Qy 3541 GTCCCTTACTCCGTGCTCTTGCGGGTGTATCTGGTAACGACTCTGGGGTC 3600
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3541 GTCCCTTACTCCGTGCTCTTGCGGGTGTATCTGGTAACGACTCTGGGGTC 3600

Qy 3601 ATTGGTTGAGGTCAACAAATCCGATGACCCCTATTGGGTGTGTGTGCGCGGGGCCCCAT 3660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3601 ATTGGTTGAGGTCAACAAATCCGATGACCCCTATTGGGTGTGTGTGCGCGGGGCCCCAT 3660

Qy 3661 GGCTGTTGCCAAGGGTTCTTCAGGTGCCCCGATTCTGCTGCCTCCTCCCGGGGCATGTTATTGG 3720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3661 GGCTGTTGCCAAGGGTTCTTCAGGTGCCCCGATTCTGCTGCCTCCTCCCGGGGCATGTTATTGG 3720


Qy 3721 GATGTTCACCGCTGCTAGAGAAATTCTGCGGTTCAGTCGGTCAGTCAGATTAGGGTTAGGCCGTT 3780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3721 GATGTTCACCGCTGCTAGAGAAATTCTGGCGGTTCAGTCGGTCAGTCAGATTAGGGTTAGGCCGTT 3780

Qy 3781 GGTGTGTGCTGATACCAGTCCCAGTACACACAGCCACTCTTGATACAAAACCTAC 3840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3781 GGTGTGTGCTGATACCATCCCAGTACACACAGCCACTCTTGATACAAAACCTAC 3840

Qy 3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCACTGGCAGGGCAAGTCAAC 3900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCACTGGCAGCGCAAGTCAAC 3900

Qy 3901 CAAATTACCCACTTCTTACATGCAGGAGAAGTATGAGTCTTGGTCCTAAATCCCAGTGT 3960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3901 CAAATTACCCACTTCTTACATGCAGGAGAAGTATGAGTCTTGGTCCTAAATCCCAGTGT 3960

Qy 3961 GGCTACAACAGCATCAATGCCAAAGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGGCAT 4020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3961 GGCTACAACAGCATCAATGCCAAAGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGGCAT 4020

Qy 4021 CTATTTTAATGGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGGCAT 4080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4021 CTATTTTAATGGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGGCAT 4080

Qy 4081 GTACCTGACCGGAGCATGTTCCCGGAACTATGATGTAATCATTGTGACGAATGCCATGC 4140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4081 GTACCTGACCGGAGCATGTTCCCGGAACTATGATGTCATCATTGTGACGAATGCCATGC 4140

Qy 4141 TACCGATGCAACCACCGTGTTGGCATTGGAAAGGTCCTACCCCCCTGGAGTAATCCCCTTC 4200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4141 TACCGATGCAACCACCGTGTTGGCATTGGAAAGGTTCTACCCCCCTGGAGTAATCCCCTTC 4200

Qy 4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTTGGCGTGTGGGGGACAC 4260
            |||||||||||||||||||||||||||||||||||||||||||||
Db 4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTTGGCGTGTGGGGGACAC 4260

Qy 4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTCATGGAAAAAGAT 4320
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTCATGGAAAAAGAT 4320

Qy 4321 TAAGGAGAGAAAATCTGAAGAAACTGGGAGGGAGACACCTTATCTTTGAGGCTACCAAACACTG 4380
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4321 TAAGGAGAGAAAATCTGAAGAAACTGGGAGGGAGACACCTTATCTTTGAGGCTACCAAACACTG 4380

Qy 4381 TGATGAGCTTGCTAACGAGTTAGCTCGAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4381 TGATGAGCTTGCTAACGAGTTAGCTCGAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440

Qy 4441 ATGTGACATCTCAAAAATCCCTGAGGGCGACTTGTTGCCACTGATGCCTTGTG 4500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4441 ATGTGACATCTCAAAAATCCCTGAGGGCGACTTGTTGCCACTGATGCCTTGTG 4500

Qy 4501 TACAGGGTACACTGGTGACTTGTATGACTGGTAGGTAGAAGGCAC 4560
            |||||||||||||||||||||||||||||||||||||||||||||
Db 4501 TACAGGGTACACTGGTGACTTGTATGACTGGTAGGTAGAAGGCAC 4560

Qy 4561 ATGCCATGTTGACCTTGACCCTACTTTCGTTGTGTGTGGGGTTTCAGC 4620
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db 4561 ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTGTCGTTGTGTGTGGGGTCTCAGC 4620

Qy 4621 AATAGTTAAAGGCCAGGCCGTAGGGGCCCGGTCGGGGTCTGGCATATACTACTA 4680
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4621 AATAGTTAAAGGCCAGGCCGTAGGGGCCCGGTCGGGGTCTGGCATATACTACTA 4680

Qy 4681 TGTGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4681 TGTGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCCTT 4740

Qy 4741 CGACGCGCCAAGGCCATGGTATGGTTTGTCATCAACAGCTCAAACTATTCTGGACAC 4800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4741 CGACGCGCCAAGGCCATGGTATGGTTTGTCATCAACAGCTCAAACTATTCTGGACAC 4800

Qy 4801 CTATCGCGACCCAACCTGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT 4860
Db 4801 
```

```
Right column:

Db  5881  TATCCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG  5940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  5941  CAGGAAGATACTGGGCATTCGAGGCATCTACCCCTGGAGTGTCATATCAGCTTGCAT    6000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941  CAGGAAGATACTGGGCATTCGAGGCATCTACCCCTGGAGTGTCATATCAGCTTGCAT    6000

Qy  6001  CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGCCTCATTGCTTGGGGTCTAGAGAT  6060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001  CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGCCTCATTGCTTGGGGTCTAGAGAT  6060

Qy  6061  TTGGCAGTATGTGTGCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGCTGGAGTTCA  6120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061  TTGGCAGTATGTGTGCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGCTGGAGTTCA  6120

Qy  6121  GAGCATGGTAACATTCCTGGTTGTCCTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC    6180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121  GAGCATGGTAACATTCCTGGTTGTCCTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC    6180

Qy  6181  CTGGATTGGATCAGTATGCTCCAAGCACGCTGTCCATGCGGTGCTGAACTCATCTTTTC   6240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181  CTGGATTGGATCAGTATGCTCCAAGCACGCTGTCCATGCGGTGCTGAACTCATCTTTTC   6240

Qy  6241  TGTTGAGAATGGTTTTGCAAAACTTTACAAAGGACCCAGAAACTTGTTCAAATTACTGGAG  6300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241  TGTTGAGAATGGTTTTGCAAAACTTTACAAAGGACCCAGAAACTTGTTCAAATTACTGGAG  6300

Qy  6301  AGGGGCTGTTCCAGTCAACGCTAGGCTGTGTGGGTCGGCTAGACCTGGAGCCCAA       6360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301  AGGGGCTGTTCCAGTCAACGCTAGGCTGTGTGGGTCGGCTAGACCTGGAGCCCAA       6360

Qy  6361  GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGGAGA  6420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361  GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGGAGA  6420

Qy  6421  TCACATTTTTGTTACAGCAGTATCCTTCCAAATGTCTGTTTCACCCAGTGCCCCCAAC    6480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421  TCACATTTTTGTTACAGCAGTATCCTTCCAAATGTCTGTTTCACCCAGTGCCCCCAAC    6480

Qy  6481  CTTGAGAGCTGCAGTCGAGCGTCCTGAGGTTCAGTGTTATCTAGGTGAGCCCCAA       6540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481  CTTGAGAGCTGCAGTCGAGCGTCCTGAGGTTCAGTGTTATCTAGGTGAGCCCCAA       6540

Qy  6541  AACTCCTTGACGACATCGCTTGCTGCGGTAAGGGTAAAACTGTTAA              6600
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  6541  AACTCCTTGACGACATCGCTTGCTGCGGTAAGGGTAAAACTGTTAA              6600

Qy  6601  GCTTCCCTTGCCGTGTGACGGTGCGCATGCGCACACACCTGGTTGCGCCATGCGAGTGA   6660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601  GCTTCCCTTGCCGTGTGACGGTGCGCATGCGCACACACCTGGTTGCGCCATGCGAGTGA   6660

Qy  6661  TGCACTTGACAAATGACTGTATTCCACAAACACTCCTAGTGATGAAGCCGCAGT      6720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661  TGCACTTGACAAATGACTGTATTCCATAAACACTCCTAGTGATGAAGCCGCAGT      6720

Qy  6721  GTCCGCTCTTGTTTTCAAACAGGAGTTGCGGCGGCTACAAACCAATTGCTTGAGGCAATTTC  6780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721  GTCCGCTCTTGTTTTCAAACAGGAGTTGCGGCGGCTACAAACCAATTGCTTGAGGCAATTTC  6780

Qy  6781  AGCTGGCGTTGACACCACCAAACTGCCAGCCCCTCCATCCGAAGAGGTAGTGGTAAGAAA   6840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781  AGCTGGCGTTGACACCACCAAACTGCCAGCCCCTCCATCCGAAGAGGTAGTGGTAAGAAA   6840

Qy  6841  GCGGCCAGTTCCGGGCAAGAACTGGTTCGCTTCGGTCTCTTGTCAGC            6900
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841  GCGGCCAGTTCCGGGCAAGAACTGGTTCGCTTCGGTCTCTTGTCAGC            6900

Qy  6901  AGGAGTGTCATGTCCTGAAAGCGTGCAACGAAGTGACCCGTTAGAAGGTCCTTCAAACCT   6960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901  AGGAGTGTCATGTCCTGAAAGCGTGCAACGAAGTGACCCGTTAGAAGGTCCTTCAAACCT   6960

Qy  6961  CCCTTCTTCACCACTTGTTCTACAGTTGGCCATGCCGATGCCGATGCCCGGGTGA       7020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961  CCCTTCTTCACCACTTGTTCTACAGTTGGCCATGCCGATGCCGATGCCCGGGTGA       7020
```

```
Left column:

Db  4801  CTATCCGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT  4860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4861  CTTTTCTATGGTCAACCCCGAACCTTCATTTGTCAATACTGCAAAAAGAGAACTGCTGACAA  4920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861  CTTTTCTATGGTCAACCCCGAACCTTCATTTGTCAATACTGCAAAAAGAGAACTGCTGACAA  4920

Qy  4921  TTATGTTTGTTGACTGAGCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCC     4980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921  TTATGTTTGTTGACTGAGCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCC     4980

Qy  4981  CAATGACGCACCACGGTGGCAGGAGCCCGGCTTGGGGAAAAACCTTGTGGGGTTCTGTG    5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981  CAATGACGCACCACGGTGGCAGGAGCCCGGCTTGGGGAAAAACCTTGTGGGGTTCTGTG    5040

Qy  5041  GCGCTTGGACGCGGCCTGACGCGCCTGTCCTGGCCCAGAGCCAGCGAGGTGACCAGATACCA  5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041  GCGCTTGGACGCGGCCTGACGCGCCTGTCCTGGCCCAGAGCCAGCGAGGTGACCAGATACCA  5100

Qy  5101  AATGTGCTTCACTGAAGTCAATACTTCTGGACAGCGACTCGCTGTTGGGCGTTGGAGT    5160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101  AATGTGCTTCACTGAAGTCAATACTTCTGGACAGCGACTCGCTGTTGGGCGTTGGAGT    5160

Qy  5161  GGCTATGGCTTATCTAGCCATTGACACTTTGGCGCCACTTGCTGTGCGGCGGTTGCTCGTC  5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161  GGCTATGGCTTATCTAGCCATTGACACTTTGGCGCCACTTGTGTGCGGCGGTTGCTCGTC  5220

Qy  5221  TATTACATCAGTCCCTACCGGTGCTACTGTCGCCCCCAGTGGTTGACGAAGAAGAAATCGT  5280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221  TATTACATCAGTCCCTACCGGTGCTACTGTCGCCCCCAGTGGTTGACGAAGAAGAAATCGT  5280

Qy  5281  GGAGGAGTGTGCATCATTCCCTTGACGGGCCATGGTTGCTGCAATTGACAAGCTGAA     5340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281  GGAGGAGTGTGCATCATTCCCTTGACGGGCCATGGTTGCTGCAATTGACAAGCTGAA     5340

Qy  5341  GAGTACAATCACCACAGTGCTTCACACTAGTCCTTTCAAATTGGAAAACCGCCCTTGAAAAACTTAACAC  5400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341  GAGTACAATCACCACAGTGCTTCACACTAGTCCTTTCAAATTGGAAAACCGCCCTTGAAAAACTTAACAC  5400

Qy  5401  CTTTCTTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGTTTAGT   5460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401  CTTTCTTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGTTTAGT   5460

Qy  5461  CACTTTACCTTGACAATCGGTGGGTTTGTCTTTGCTTGTCATTGCGGGTATTACTAC     5520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461  CACTTTACCTTGACAATCGGTGGGTTTGTCTTTGCTTGTCATTGCGGGTATTACTAC     5520

Qy  5521  CCCACTACCTCACAAGATCAAAATGTTCCGTGTCATTATTTGGAGGCGCAATTGCGTCCAA  5580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521  CCCACTACCTCACAAGATCAAAATGTTCCGTGTCATTATTTGGAGGCGCAATTGCGTCCAA  5580

Qy  5581  GCTTACAGACGCTAGAGCGCACTGGGCGTTCAATCGGCCGGCGGCTGCGCGGGAACAGCTCT  5640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581  GCTTACAGACGCTAGAGCGCACTGGGCGTTCAATCGGCCGGCGGCTGCGCGGGAACAGCTCT  5640

Qy  5641  TGGTACATGGACATCGGTGGGTTTGTCTTTGTGTTTGTCTTTCATTGCGCATGCTCGCCGCTC  5700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641  TGGTACATGGACATCGGTGGGTTTGTCTTTGTGTTTGTCTTTCATTGCGCATGCTCGCCGCTC  5700

Qy  5701  ATCCACTGCTTGCTTTGACAATTTAAATGCTTGATGGGTGAGTGGCCCACTATGGATCAGCT  5760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701  ATCCACTGCTTGCTTTGACAATTTAAATGCTTGATGGGTGAGTGGCCCACTATGGATCAGCT  5760

Qy  5761  TGCTGGTTTAGTCTACTCCGGCGTTCAATCCGGCCGGCGTCTCTTGTCAGC            5820
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761  TGCTGGTTTAGTCTACTCCGGCGTTCAATCCGGCCGGCGTCTCTTGTCAGC            5820

Qy  5821  TTGTGCAATGTTTGCTTTGACAACAGCAGGGCCGGCCAACACTCCTTAC             5880
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821  TTGTGCAATGTTTGCTTTGACAACAGCAGGGCCGGCCAACACTCCTTAC             5880

Qy  5881  TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG  5940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy  7021  GTGTAACCCTTCACTGCAATTGGATGTGCAATGACCGAAACAGGCGGAGCCCTGATGA  7080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021  GTGTAACCCTTCACTGCAATTGGATGTGCAATGACCGAAACAGGCGGAGCCCTGATGA  7080

Qy  7081  TTTACCCAGTTACCCTCCAAAAAGGAGGTCTCGATGGTCAGAGCAAAGTTGGTCGAC  7140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081  TTTACCCAGTTACCCTCCAAAAAGGAGGTCTCGATGGTCAGAGCAAAGTTGGTCAAC  7140

Qy  7141  GGCTACAACCGCTTCCAGCTACGTTACTGGCCCCCGTACCCTAAGATACGGGAAAGGA  7200
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141  GACTACAACCGCTTCCAGCTACGTTACTGGCCCCCGTACCCTAAGATACGGGAAAGGA  7200

Qy  7201  TTCCACTCAGTCAGCCCCGCCAAACGGCCTACAAAAAGGAGTTGGGAAAGAGTGAGTT  7260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201  TTCCACTCAGTCAGCCCCGCCAAACGGCCTACAAAAAGGAGTTGGGAAAGAGTGAGTT  7260

Qy  7261  TTCGTGCAGAGCTACACCTGGACCACGTGACCTGATTAGCTTCAAAAACTGCTTCTAAAGT  7320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261  TTCGTGCAGAGCTACACCTGGACCACGTGACCTGATTAGCTTCAAAAACTGCTTCTAAAGT  7320

Qy  7321  TCTGTCTGCAACTCGGCCATCACTACTAGTGGTTTCCTCAAACAAAGATCATTGGTGTATGT  7380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7321  TCTGTCTGCAACTCGGCCATCACTACTAGTGGTTTCCTCAAACAAAGATCATTGGTGTATGT  7380

Qy  7381  GACTGAGCGCCGGGATGCGGGAGCTTAGAAGCAAAACAAAAAGTCACTATTAATAGACAACCTCT  7440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381  GACTGAGCGCCGGGATGCGGGAGCTTAGAAGCAAAACAAAAAGTCACTATTAATAGACAACCTCT  7440

Qy  7441  GTTCCCCCATCATACCACAAGCAAGCAGATTGGCTAAGGAAAAGCTTCAAAAGTGT  7500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441  GTTCCCCCATCATACCACAAGCAAGCAGATTGGCTAAGGAAAAGCTTCAAAAGTGT  7500

Qy  7501  CGGGTCATGTGGGACTATGAGAGTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTC  7560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7501  CGGGTCATGTGGGACTATGAGAGTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTC  7560

Qy  7561  CCACATCACTGGCCTTCGTTCGGGGCACTGCCTTCGTTCTGGAGCAGCCCCAAGGCTGTTCT  7620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7561  CCACATCACTGGCCTTCGTTCGGGGCACTGCCTTCGTTCTGGAGCAGCCCCAAGGCTGTTCT  7620

Qy  7621  GGACTTGCAGAAGTGTGTGAGGCAGTGAGATGCCGAGTCATTATCGGCAAACTGTGAT  7680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7621  GGACTTGCAGAAGTGTGTGAGGCAGTGAGATGCCGAGTCATTATCGGCAAACTGTGAT  7680

Qy  7681  AGTTCCAAAGGAGGAGGTCTTCGTGAAGACAACCAACAAAGAAACCCCCAAG  7740
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7681  AGTTCCAAAGGAGGAGGTCTTCGTGAAGACAACCAACAAAGAAACCCCCAAG  7740

Qy  7741  GCTTATCTCGTACCCCACCTTGAGATGAGTGTTGAGAGAATGTGTTGAGAGATGCGTACGGT  7800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7741  GCTTATCTCGTACCCCACCTTGAGATGAGTGTTGAGAGAATGTGTTGAGAGATGCGTACGGT  7800

Qy  7801  TGCTCCTGACGTAGTTAAAGTGTCATTGGAGGATGCGTACGGGTTTGTGAGATCCAGTAC  7860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7801  TGCTCCTGACGTAGTTAAAGTGTCATTGGAGGATGCGTACGGGTTTGTGTCGACCCACGTAC  7860

Qy  7861  CCGGTGTCAAGCGTCTGTTGTCGATGTGGTCACCCGATGCGAGCGCCAATGCGGATAC  7920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7861  CCGGTGTCAAGCGTCTGTTGTCGATGTGGTCACCCGATGCAGCGCCAATGCGGATAC  7920

Qy  7921  AGTGTGTTTGACAGTTGTACCATCATCACACCCGGAGATATCATGTGGAGACAGACATCTACTC  7980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7921  AGTGTGTTTGACAGTTGTACCATCATCACACCCGGAGATATCATGTGGAGACAGACATCTACTC  7980

Qy  7981  AGCAGCTAAACTCAGTGACCAACACCGGAGTAAACTCAGTGACCAACACCGCCGATATCGTAGGTGTAGGTGTAGGTC  8040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7981  AGCAGCTAAACTCAGTGACCAACACCAGCTGCCGATATCGTAGGTGTAGGTGTAGGTC  8040

Qy  8041  CGCTGGAGGAGGACCGGATGATGCGCTTATGATGGCCGAGAATCGGATATCGTTAGGTGTAGGTC  8100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8041  CGCTGGAGGAGGACCGGATGATGCGCTTATGATGGCCGAGAATCGGATATCGTTAGGTGTAGGTC  8100
```

```
Qy  8101  TTCCGGGGTCTATACTACCTCAAGTTCCAACAGTTTGACCTGCTGGCTGAAGGTAAATGC  8160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8101  TTCCGGGGTCTATACTACCTCAAGTTCCAACAGTTTGACCTGCTGGCTGAAGGTAAATGC  8160

Qy  8161  TGCAGCCAACAGGCTGGCATGAGAAGAACCCTGCCTTCCTTATTTGCGGCGATGATTGCAC  8220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161  TGCAGCCAACAGGCTGGCATGAGAAGAACCCTGCCTTCCTTATTTGCGGCGATGATTGCAC  8220

Qy  8221  CGTAATTTGGAAGAGCGCCGGAGCAGATGCAGACAAGCAATGCGTGTCTTTGCTAG  8280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8221  CGTAATTTGGAAGAGCGCCGGAGCAGATGCAGACAAGCAAATGCGTGTCTTTGCTAG  8280

Qy  8281  CTGGATGAAGGTGATGGGTGCACCCAAGATGTGTGCCTCAACCCAAATACAGTTTGGA  8340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281  CTGGATGAAGGTGATGGGTGCACCCAAGATGTGTGCCTCAACCCAAATACAGTTTGGA  8340

Qy  8341  AGAATTAACAACATGCTCATCATCAAAAGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA  8400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8341  AGAATTAACAACATGCTCATCATCAAAAGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA  8400

Qy  8401  CTACTTCTTACAAGATCTCGTATCCCCCTTGGCAGGTGCTCTCGCGAGGGTCTGGG  8460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8401  CTACTTCTTACAAGATCTCGTATCCCCCTTGGCAGGTGCTCTCGCGAGGGTCTGGG  8460

Qy  8461  ATACAACCCCAGTGCTGCGTGGATTGGGTATCTAATACATCAATACACTCACTTGTGTGGGT  8520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8461  ATACAACCCCAGTGCTGCGTGGATTGGGTATCTAATACATCAATACACTCACTTGTGTGGGT  8520

Qy  8521  TAGCCGTGTGTGGCGTCGTCCATTTCACTTCAGGAGCAGATGCTCTTTGAGGACAAAACTTCCCGA  8580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8521  TAGCCGTGTGTGGCGTCGTCCATTTCACTTCAGGAGCAGATGCTCTTTGAGGACAAAACTTCCCGA  8580

Qy  8581  GACTGTGACCTTTGACTGTGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG  8640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8581  GACTGTGACCTTTGACTGTGTATGGGAAAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG  8640

Qy  8641  CATCATTGCTGTGTGTGGCACGGTATTGGAGCTTTCTCCGGTGCGCCTACACCAACGCTGA  8700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8641  CATCATTGCTGTGTGTGGCACGGTATTGGAGCTTTCTCCGGTGCGCCTACACCAACGCTGA  8700

Qy  8701  GATCCTCAGAGTTTCCCAATCCAATCACTAACAGACAGACATGACCAGCGCCCCTGCGAGCCTGGCG  8760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8701  GATCCTCAGAGTTTCCCAATCCAATCACTAACAGACAGACATGACCAGCGCCCCCTGCGAGCCTGGCG  8760

Qy  8761  AAAGAAGCCAGGGCGGTCCTCCGCCAAGAGGGCGTGGGCGAGCACCGCAAAATT  8820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8761  AAAGAAGCCAGGGCGGTCCTCCGCCAAGAGGGCGTGGGCGAGCACCGCCAAAATT  8820

Qy  8821  GGCTCGCTTCCTTCTCTCGGCGATGCCATCTCTACAGATTTGGATAAGAACGAG  8880
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8821  GGCTCGCTTCCTTCTCTCGGCGATGCCATCTCTACAGATTTGGATAAGAGCGAG  8880

Qy  8881  CGTGGCTCGGTACACCACTTTCAATTATTGTGATGTGTTTACTCCCCGGGATGTGTT  8940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8881  CGTGGCTCGGTACACCACTTTCAATTATTGTGATGTTTTTACTCCCCGGGATGTGTT  8940

Qy  8941  TATTACACCAGAGGAAGATTGCAGAAGTTCCTTGTGAAGTATTTGGGCTGTCATTGTTTT  9000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8941  TGTTACACCAGAGGAAGATTGCAGAAGTTCCTTGTGAAGTATTTGGGCTGTCATTGTTTT  9000

Qy  9001  TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATCAGTGAACCCCAAATTCAAAATTAA  9060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9001  TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATCAGTGAACCCCAAATTCAAAATTAA  9060

Qy  9061  CTAACAG----TTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCAACAGGGGAGACCCC  9116
          |        ||||||||||||||||||||||||||||||||||||||||||
Db  9061  TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCAACAGGGGAGACCCC  9120

Qy  9117  GGGCTTAACGACCCCGC  9133
          |||||||||||||||||
Db  9121  GGGCTTAACGACCCCGC  9137
```

```
RESULT 10
AAA55280
ID    AAA55280   standard; DNA; 8912 BP.
XX
AC    AAA55280;
XX
DT    06-AUG-2003  (revised)
DT    30-AUG-2000  (first entry)
XX
DE    Hepatitis GB virus B nucleotide sequence SEQ ID NO:11.
XX
KW    Hepatitis GB virus; HGBV; diagnosis; therapeutic; immunogenic; infection;
KW    detection; characterisation; hepatitis; ss.
XX
OS    Hepatitis GB virus.
XX
PN    US6051374-A.
XX
PD    18-APR-2000.
XX
PF    07-JUN-1995;  95US-00488445.
XX
PR    14-FEB-1994;  94US-00196030.
PR    13-MAY-1994;  94US-00242654.
PR    29-JUL-1994;  94US-00283314.
PR    23-NOV-1994;  94US-00344185.
PR    23-NOV-1994;  94US-00344190.
PR    30-JAN-1995;  95US-00377557.
XX
PA    (ABBO ) ABBOTT LAB.
XX
PI    Dawson GJ, Leary TP, Muerhoff AS, Pilot-Matias TJ, Buijk SL;
PI    Mushahwar IK, Simons JN, Desai SM, Erker JC, Schlauder GG;
XX
DR    WPI; 2000-338307/29.
XX
PT    Detecting target hepatitis GB virus nucleic acid in a test sample
PT    suspected of containing HGBV comprises reacting the test sample the HGBV
PT    polynucleotide probe and detecting the complex that contains target HGBV.
XX
PS    Example 4; Col 131-140; 369pp; English.
XX
CC    The present invention describe a method for detecting target hepatitis GB
CC    virus (HGBV) nucleic acid (THN) in a test sample (T) suspected of
CC    containing HGBV. The method involves reacting (T) with a HGBV
CC    polynucleotide probe (I) containing 15 contiguous nucleotides, and which
CC    selectively hybridises to the HGBV genome or its full complement, and
CC    detecting the complex that contains THN, indicating the presence of
CC    target HGBV. The method is used for detecting target HGBV nucleic acid in
CC    the test sample suspected of containing HGBV and for characterisation of
CC    newly ascertained etiological agent of non-A, non-B, non-C, non-D and non
CC    -E hepatitis causing agents collectively termed as hepatitis GB virus.
CC    protein sequences used in the exemplification of the present invention.
CC    AAA55270 to AAA55489 and AAB08985 to AAB09480 represent nucleotide and
CC    (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ    Sequence 8912 BP; 2019 A; 2169 C; 2265 G; 2400 T; 0 U; 59 Other;

Query Match             92.5%;  Score 8692.6;  DB 3;  Length 8912;
Best Local Similarity   98.7%;  Pred. No. 0;
Matches 8781;  Conservative   56;  Mismatches   52;  Indels   10;  Gaps   8;
```

Qy   196 TGGGTTCGGTGGTGGTGCGGCTTTAGGCAGCCTCACCACCTCCCAGCCACCACCTCCCAGATAGAGC 255
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11 TGGGTTCGGTGGTGGTGCGGCTTTAGGCAGCCTCACCACCTCCCAGCCACCACCTCCCAGATAGAGC 70

Qy   256 GGCGGCACTGTAGGGAAGACCGGGACCGGGTCACTACCAAGGACGCAGACCTCTTTTTGA 315
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71 GGCGGCACTGTAGGGAAGACCGGGACCGGGTCACTACCAAGGACGCAGACCTCTTTTTGA 130

Qy   316 GTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTTGTTGGGATGGTTGGGGT 375
              |||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db   131 GTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACCTAYATGTTGTTGGGATGGTTGGGGT 190

Qy   376 TAGGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGATCTGGGAGTCTTGG 435
              ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db   191 TAGGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGATCTGGGAGTCTCGTAG 250

Qy   436 ACGTAGCACATGCCGTGTTATTTCTACTCAAACAAGTCCTGTACCTGCCGCCCCAGAACGCG 495
              ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db   251 ACCGTAGCACATGCCTGTTATTTCTCACTCAAACAAGTCCTGTACCTGCGCCCCAGAACGCG 310

Qy   496 CAAGAACAAGCAGACGCAGCTTCATATCTGTGTCCATTAAAACATCTGTTGAAAGGGG 555
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   311 CAAGAACAAGCAGACGCAGCTTCATATCTGTGTCCATTAAAACATCTGTTGAAAGGGG 370

Qy   556 ACAACGAGCAAAGCGCAAAGTCCAGCGCGATGTCGGCCTCGGCGGCCTGTAATTACAAAATTGCTGG 615
              |||||||||||||:|||||||||||||||||||||||||||||:||||||||||||||||||||||
Db   371 ACAACGAGCAARGCGCAAAGTCCAGCGCGATGTCGGCCTCGGCCTCGTAATTACAAAATTGCTGG 430

Qy   616 TATCCATGATGGCTTGCGACACATTGGCTCAGGCTGCTTGCCAGCTCATGGTTGGGGACG 675
              |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db   431 TATCCATGATGGCTTGCGACACATTGGCTCAGGCTGCTTRCCAGCTCATGGTTGGGGACG 490

Qy   676 CCAAGACCCTGCCATAAGTCTCGCAATCTTGGAATCTTCTGGATTACCCTTTGGGGTG 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   491 CCAAGACCCTCGCCATAAGTCTCGCAATCTTGGAATCTTCTGGATTACCCTTTGGGGTG 550

Qy   736 GATTGGTGATGTTACAACTCACACACCTCTAGTAGGCCGCTGGTGGCAGGAGGGGTCGT 795
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   551 GATTGGTGATGTTACAACTCACACACCTCTAGTAGGCCCGTGGTGGCAGGAGCGGTCGT 610

Qy   796 TCGACCAGTCTGCCAGATAGTAGCGCTTGCTGGAGGATGGAGTCAACTGGGCTACTGGTTG 855
              |||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db   611 TCGACCAGTCTGCCAGATAGTAGCGCTTGCTGGAGGATGGAGTCAACTGGGCTACTGGTTG 670

Qy   856 GTTCGGTGTCCACCTTTTGTGGTATGTCGCTATCTTGGCCTCGTCCCTGTAGTGGGGC 915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   671 GTTCGGTGTCCACCTTTTGTGGTATGTCGCTATYTTGGCCTGTCCCTGTAGTGGGGC 730

Qy   916 GCGGGTCACTGACCCAGACACAAATACCACACACATCCTGCCAGCAATTGCTGCCAGCGTAATCA 975
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   731 GCGGGTCACTGACCCAGACACAAATACCACAATACCTGACCAATTGCTGCCAGCGTAATCA 790

Qy   976 GGTTATCTATTGTTCTCCTTCCACTTGCCTACACGAGCCTGGTTGTGTGTGATCGTCGTCGGA 1035
              ||||||:||||||||||||||||::|||||||||||||||||||||||||||||||||||:|
Db   791 GGTTATCTAYTGTTCTCCTTCCACTTGCCTACACGAGCCTGGTTGTGTGTGATCGTCGTGYGGA 850

Qy   1036 CGAGTGCTGGGTTCCCGCCAATCGTACATCTCACACCCTTCACACCCTTCCAATTGGACTGGCACGGA 1095
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   851 CGAGTGCTGGGTTCCCGCCAATCCGTACATCCACACCCTTCACACCCTTCCAATTGGACTGGCACGGA 910

Qy   1096 CTCCTTCTTGGCTGACCACCATTGATTTTGTTATGGGCGTCTTGTGACCTGTGACGCCCT 1155
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db   911 CTCCTTCTTGGCTGACCACCATTGATTTTGTTATGGGCGTCTTGTGACCTGTGACGCCCT 970

Qy   1156 TGACATTGGTGAGTTGTGGTGCGTGTGTATTAGTCGGTGACTGGCGTTGTCAGGCACTG 1215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   971 TGACATTGGTGAGTTGTGGTGCGTGTGTATTAGTCGGTGACTGGCGTTGTCAGGCACTG 1030

Qy   1216 GCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGAAGTGCCCACTGGAAT 1275
              ||||||||||||||||||||||||||||||:||||||||||||||:||||||||||||||
Db   1031 GCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGAAKTGCCCACTGGAAT 1090

Qy   1276 AGATCCTGGGTTCCTAGGGTTTATCGGGGTGATGGCCGCAAGGTCGAGGCTGTCATCTT 1335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1091 AGATCCTGGGTTCCTAGGGTTTATCGGGGTGATGGCCGCAAGGTCGAGGCTGTCATCTT 1150

Qy   1336 CTTGACCAAACTGCTTCACAAGTACCATACGCTATTGCGACTATGTTTAGCAGTGTACA 1395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1151 CTTGACCAAACTGCTTCACAAGTACCATACGCTATTGCGACTATGTTTAGCAGTGTACA 1210

Qy   1396 CTACCTGGGGTTGGCGCTCTGATCTACATATGCCTCTCGGGGCAAGTGTATCAGTTGCT 1455
              |||||||||||||||||||||||||||||||::|||||||||||||||||||||||||
Db   1211 CTACCTGGGGTTGGCGCTCTGATCTACATACGYGCCTCTCGGGGCAAGTGTATCAGTTGCT 1270

```
Qy   1456  CCTAGGGCCTTATGCTTTACATAGAAGCGACCTCTGAAACCCATCAGGGTGCCCACTGG  1515
           |||||||||||||  |||||||||||||||||||||||||||||||||  |||||||||||
Db   1271  CCTAGGGCCTTAYGCTTTACATAGAAGCGACCTCTGAAACCCYATCAGGGTGCCCACTGG  1330

Qy   1516  ATGCTCAATAGCTGAGTTTTGCTCGCCTTTGATGATACCATGTCCTTGCCACTCTTATTT  1575
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1331  ATGCTCAATAGCTGAGTTTTGCTCGCCTTTGATGATACCATGTCCTTGCCACTCTTATTT  1390

Qy   1576  GAGTGAGAATGTGTCAGAAGTCATTTGTTACAGTCCAAAGTGACCAGGCCTATCACTCT  1635
           ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   1391  GAGTGAGAATGTGTCAGAAGTCATTTGTTACAGTCGACCAAAGTGACCAGGCCTGTCACTCT  1450

Qy   1636  AGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGGTGCGAGGGGATGTAT  1695
           |||||||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1451  AGAGTATAABAACTCCATATCTTGGTACCCCTATACAATCCCTGGTGCGAGGGGATGTAT  1510

Qy   1696  GGTTAAATTCAAAAATAACACATGGGGTGCTGCCATTCGCAATGTGCCATCGTACTG  1755
           ||||||||||||||||||||||||||||||   ::|||||||||||||||||||||||
Db   1511  GGTTAAATTCAAAAATAACACATGGGGTGCTGCGCG-WWTCGCAATGTGCCATCGTACTG  1569

Qy   1756  CACTATGGGCACTGATGCAGTGTGGAACGACACTCGCAACTTACGAAGCATGCGGTGT  1815
           |||||||||||||||||||||||||| :|||||||||||||||||||||||||||||
Db   1570  CACTATGGGCACTGATGCAGTGTGGAASSACAGTCGCAACACTTACGAAGACACATGCCGGTGT  1629

Qy   1816  ATACCCTGGGTCTAAAAGAAAATGTTAAACCTCATAATTGGATGTCAGGCCATTGTATT  1875
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1630  ATACCCTGGGTCTAAAAGAAAATGTTAAACCTCATAATTGGATGTCAGGCCATTGTATT  1689

Qy   1936  TGAGGGATCAGATACCCTATAGTTTATACTTTTTATGACCCTGTGAATTCCACTCTCCACC  1995
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   1750  TGAGGGATCAGATACCCTATAGTTTATACTTTTTATGACCCTGTGAATTCCACTCTCCACC  1809

Qy   1996  ACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCTGTGGTACGTTCTTGGTTACA  2055
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1810  ACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCTGTGGTACGTTCTTGGTTACA  1869

Qy   2056  GGTTCCGCAAGGGTTTTACAGTGATGTGAAAGAAGAAGACCTAGCCACAGAGGATTGATC  2115
           ||||||||   ::|||||||||||||||||||||||||||||||||||||||||||||||
Db   1870  GGTTCCGCAA-GGTTTTACAGTGATGTGAAAGAAGAAGACCTAGCCACAGAGGATTGATC  1928

Qy   2116  CAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGGCCACGGGGTGCTTTGTCTCTTACGGG  2175
           |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   1929  CAAAGCCTGGAAAAATTATCAGYTCTTATATTCCGGCCACGGGGTGCTTTGTCTCTTACGGG  1988

Qy   2176  AGTTACCACCAAGGCCGTGGTGTCTGCTAATTCTTCATCTTGCTATCTCCGGCTGTCGTGGC  2235
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1989  AGTTACCACCAAGGCCGTGGTGTCTGCTAATTCTTCATCTTGCTATCTCCGGCTGTCGTGGC  2048

Qy   2236  TTTAGCTACCTCTGTTACTTGTCCCTTGTTTTGGGGCGCTTCTGGTTACCCTTTGCG  2295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2049  TTTAGCTACCTCTGTTACTTGTCCCTTGTTTTGGGGCGCTTCTGGTTACMCTTTGCG  2108

Qy   2296  TCCCTGTGTCTCCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGTTTTGTCTAAAGCTCA  2355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2109  TCCCTGTGTCTCCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGTTTTGTCTAAAGCTCA  2168

Qy   2356  AGTAGCTTCCTTTTGCTTTGATTTCTTCATCTGTTGCTATCTCCGCTGCTATCTCCGCCAGGCTACGTTA  2415
           |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   2169  AGTAGCTCMTTTTGCTTTGATTTCTTCATCTGTTGCTATCTCCGCTGCTATCTCCGCCAGGCTACGTTA  2228

Qy   2416  TGCTGCCCCTTTTAGGGTTTGTGTCTCGCGGCTGCGTGGTGGGTGGGCGCTTGCCCCTTAACTTTTCTTGTGC  2475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2229  TGCTGCCCCTTTTAGGGTTTGTGTCTCGCGGCTGCGTGGTGGGTGGGCGCTTGCCCCTTAACTTTCTTTGTGC  2288

Qy   2476  AGCAGCTGCTGCCCAACCAGATTATGACTGCTGGTGCGTCGACTGCTAGTGGCAGGGTTAGT  2535
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2289  AGCAGCTGCTGCCCAACCAGATTATGCTGGTGCGTCGACTGCTAGTGGCAGGGTTAGT  2348

Qy   2536  TTTGTGTGGGCCGGCCGCTAACCGTGGTCACCGGCATAGTCTCGCTTGTTGTAGGTCCTTGGCCCTCT  2595
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db
```

```
Db   2349  TTTGTGTGGGCCGGCCGACCGTGGTCA--CGCATAGCTCTGCTTGTTGTAGGTCCTTGGCCCTCT  2407

Qy   2596  GGTAGCGCTTT-TAACCCTCCTTGCATTTGGTTACGCCTGCTTCAGCTTTGATACCGAGA  2654
           |||||||||||  |||||||||::::SSTKACGCCTGCTT-AGCTTTSSTKACGCCGAGA
Db   2408  GGTAGCGCGTTTTYAACCCCTTCTTGCATTTSSTKACGCCCTGCTT-AGCTTTSSTKACGCCGAGA  2466

Qy   2655  TAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCATGTCTCGTTTGGCT  2714
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2467  TAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCATGTCTCGTTTGGCT  2526

Qy   2715  TCTTTGCTCACTTGTTACCTGCTGTGTTTAGTTAACTCCTATCTTTGGCAACGTTGGG  2774
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2527  TCTTTGCTCACTTGTTACCTGCTGTGTTTAGTTAACTCCTATCTTTGGCAACGTTGGG  2586

Qy   2775  AGAATTGGTTTGGAACGTACACTAAGAACTGCCGAGAGGTTTTCCTTGTGCTGTTTGTT  2834
           ||||||||||| ||||||||||||||||||||||||||||||||::|||||YGCGTGTTTGTT
Db   2587  AGAATTGGTTTTGGAACGTACACTAAGAACTGCCGAGAGGTTTCCTCCTTGTGYGCGTGTTTGTT  2646

Qy   2835  TCCCCGGTGCGACATCGTCTTCTTGGGACTGACCTTCGTGTGTGTGTCACGTAGCTCTTAT  2894
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2647  TCCCCGGTGCGACATCGTCTTCTTGGGACTGACCTTCGTGTGTGTGTCACGTAGCTCTTAT  2706

Qy   2895  GTTTAACATCCAGTGCGCATCGTCGTCTTAGGACTGACTCTAGGGTTAGGGCCCATAGAA  2954
           |||||||||||||||||||||||||||::|||||||||||||||||||||||||||||||
Db   2707  GTTTAACATCCAGTGCGCATCGTCGTCTTAGGACTGACTCTAGGGTTAGGGCCCATAGAA  2766

Qy   2955  TGTTTGGTGCGTCTCGGAAAGTGTCATGCTGTCTTGTATTCTCATTATGTTCTTAAGTTTTTCC  3014
           |||||||||||||||||||||||||||::|||||||||||||||||||||||||||||||
Db   2767  TGTTTGGTGCGTCTCGGAAAGTGTCATGCTGTCAYMGTTCTTGTATTCTCATTATGTTCTTAAGTTTTTCC  2826

Qy   3015  TCTTAGTGTTTGGTGAGAATGGTGTGTTTTTCTATAAGCACTTGCATGGTGATGTCTTGC  3074
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2827  TCTTAGTGTTTGGTGAGAATGGTGTGTTTTTCTAKAAGCACTTGCATGGTGATGTCTTGC  2886

Qy   3075  CTAATGATTTGCTCGAAACTACCATTTCCCTTTCCCCTTTGAAGGCAAGG  3134
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2887  CTAATGATTTGCTCGAAAACTACCATTTCCCTTTCCCCTTTGAAGGCAAGG  2946

Qy   3135  CAAGGGTCTATAGGAATGAAGAAGACGCTTGGCGTGTGGGGACACGGTTGATGGTTTGC  3194
           |||||||||||||||||||||||||::|||||||||::|||||||||||||||||||||
Db   2947  CAAGGGTCTATAGGAATGAAGAAGACGCCTTGGSKKGGSGGGACACGGTGTGATGGTTTGS  3006

Qy   3195  CCGTGTTGTGCGCTCTCGGCGACCTTGTTTCCGCAGGGTTGGCTATGCCGCCAGATGGGT  3254
           ::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3007  SCGTTGTGBGCGCTCTCGGCGACCTTGTTTCCGCAGGGTTAGCTGTTAGTCCGCCAGATGGGT  3066

Qy   3255  GGGCCATTACCGACCTTTACGGTCCTCTCTCGAGTCGTCTGAAACGTGGCACGTGTCAGCGA  3314
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3067  GGGCCATTACCGACCTTTACGGTCCTCTCTCGAGTCGTCTGAAACGTGGCACGTGTCAGCGA  3126

Qy   3315  TGGCAGTGGTCATGACTGGTCATGATAGACCCCGAACTGGACTGGAACTATCTTCAGATTAG  3374
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3127  TGGCAGTGGTCATGACTGGTCATGATAGACCCCGAACTTGACTGGAACTATCTTCAGATTAG  3186

Qy   3375  GATCTCTGGCCACTAGCTACATGGGATTGTTTGTGACAACGTGTTGTGTGACTGCTCACC  3434
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3187  GATCTCTGGCCACTAGCTACATGGGATTGTTTGTGACAACGTGTTGTGTGAATACTGCTCACC  3246

Qy   3435  ATGGCAGCAAGGGGCGCGGTTGGCCTCTATCACCCAATAACCGTTG  3494
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   3247  ATGGCAGCAACGGGGGGCCGGTTGGCCTCCATAACACCCAATAACCGTTG  3306

Qy   3495  ACGCGGCTAATGACCAGGACATCTATCAACCACCATGGAGCTGGGTCCCTTACTCGGT  3554
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3307  ACGCGGCTAATGACCAGGACATCTATCAACCACCATGGAGCTGGGTCCCTTACTCGGT  3366

Qy   3555  GCTCTTGCGGGGAGACCAAGGGTATCTGGGGTCATTGGGTTGAGGTCA  3614
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   3367  GCTCTTGCGGGGAGACCAAGGGTATCTGGGGTCATTGGGTTGAGGTCA  3426

Qy   3615  ACAAATCCGATGACCCTTATTGTGTGTGCGGGGGCCCTTCGCCAAGG  3674
           |||||||||||||||||||||||||||||||||||||||||||||||
Db
```

```
Qy  4755 CATGGTATGGTTTGTCATCAACAGAAGTCAAACTATTCTGGACACCTATCGCCACCAAC 4814
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4567 CATGGTATGGTTTGTCATCAACGAGAGTCAAACTATTCTGGACACCTATCGCCACCAAC 4626

Qy  4815 CTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCTCTTTCTATGGTCA 4874
         ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  4627 CTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCGATCTCTTCTATGGTCA 4686

Qy  4875 ACCCCGAACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAATTATGTTTGTTGA 4934
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4687 ACCCCGAACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAATTATGTTTGTTGA 4746

Qy  4935 CTGCAGCCCAACTACAACTGTGTCATCAGTATGCTATGCTCCCAATGACGCACCAC 4994
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4747 CTGCAGCCCAACTACAACTGTGTCATCAGTATGCTATGCTCCCAATGACGCACCAC 4806

Qy  4995 GGTGGCAGGGAGCCCGGCTTGGGAAAAACCTTGTGGCGCGTTGGACGGCG 5054
         ||||||||||||||||||||||||| |||||||||||||||||||||||
Db  4807 GGTGGCAGGGAGCCCGGCTTGGGAAAAAACCTTGTGGCGCGTTGGACGGCT 4866

Qy  5055 CTGACGCCTGTCTCTGGCCCAGAGCCCAGCGAGTGACCAGATACCAAATGTGCTTCACTG 5114
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4867 GTGACGCCTGTCTCTGGCCCAGAGCCCAGCGAGTGACCAGATACCAAATGTGCTTCACTG 4926

Qy  5115 AAGTCAATACTTCTGGACAGCCGCACTCGCTGTTGGTGCGCGTTGGCTTATGGCTTATC 5174
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4927 AAGTCAATACTTCTGGACAGCCGCACTCGCTGTTGGTGCGCGTTGGCTTATGGCTTATC 4986

Qy  5175 TAGCCATTGACACTTTGGCGCCACTTGTGTGCGCGGTTGCTGTCTATTACATCAGTCC 5234
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4987 TAGCCATTGACACTTTGGCGCCATTGTGTGCGCGGTTGCTGTCTATTACATCAGTCC 5046

Qy  5235 CTACCGGTGCTACTGTCGCGCCCCAGTGGTTGACGAAGAGAAATCGTGGAGGAGTGTGCAT 5294
         |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db  5047 CTACCGGTGCTACTGTCGCGCCCCAGTGTGGTTGACGAGGAGAAATCGTGGAGGAGTGTGCAT 5106

Qy  5295 CATTCATTCCCTTGGAGGCCATGTTGCTCGCAATTGACAAGCTGAAGAGTACAATCACCA 5354
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5107 CATTCATTCCCTTGGAGGCCATGTTGCTCGCAATTGACAAGCTGAAGAGTACAATCACCA 5166

Qy  5355 CAACTAGTCCTTTCACATGGAAACGCCCTTGAAAAACTAACACCTTTCTTCGGGCCTC 5414
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5167 CAACTAGTCCTTTCACATTGAAACCGCCCTTGAAAAACTAACACCTTTCTTCGGGCCTC 5226

Qy  5415 ATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTCGGTTAGTCACTTACCTGCTA 5474
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5227 ATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTCGGTTAGTCACTTACCTGCTA 5286

Qy  5475 ATCCCTTTGCATCATGCGTGTTTGCTTTCATTGCGGGGTATTACTACCCCACTACCTCACA 5534
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5287 ATCCCTTTGCATCATGCGTGTTTGCTTTCATTGCGGGGTATTACTACCCCACTACCTCACA 5346

Qy  5535 AGATCAAAATGTTCCTGTCATTATTTGGAGGCGCAATTGCGTCCAAGCTTACAGACGCTA 5594
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5347 AGATCAAAATGTTCCTGTCATTATTTGGAGGCGCAATTGCGTCCAAGCTTACAGACGCTA 5406

Qy  5595 GAGGCGCACTGGCGTTCATGATGGCCGGGAACAGCTCTTGGTACATGCGACAT 5654
         ||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db  5407 GAGRCGCACTGGCGTTCATGATGGCGYGGGAACAGCTCTTGGTACATGCGACAT 5466

Qy  5655 CGGTGGGTTTTGTCTTTGACATGCGTTAGGCGCCCTATGCTGCCGCCCTCACTCCTGCT 5714
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5467 CGGTGGGGTTTTGTCTTTGACATGCGTTAGGCGCCCTATGCTGCCGCCCTCACTGCTGCT 5526

Qy  5715 TGACATTTAAATGCTTGATGGGTGAGTGGCGCCACTACTCAGCTTGCTGGTTTAGTCT 5774
         ||||||||||||||||||||||||||:||||||||::|||||||||||||||||||||
Db  5527 TGACATTTAAATGCTTGATGGGTGAGTGGCGYCACTATGGATCAGCTTGCTGGTTTAGTCT 5586

Qy  5775 ACTTCCGCGTTCAATCCGGCGGTGTGGGCCGTCTTGTCGTGGCTTGTTGTCAATGTTTG 5834
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5587 ACTTCCGCGTTCAATCCGGCGGTGTGGGCCGTCTTGTCGTGGCTTGTTGTCAATGTTTG 5646
```

```
Db  3427 ACAAATCCGATGACCCTTATTGGTGTGTGTGCGGGGCCCTTCCCATGGCTGTTGCCAAGG 3486
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3675 GTTCTTCAGGTGCCCGATTCTGTGCTCCTCCGGGCATGTGTTATTGGGATGTTCACCGCTG 3734
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3487 GTTCTTCAGGTGCCCGATTCTGTGCTCCTCCGGGCATGTGTTATTGGGATGTTCACCGCTG 3546

Qy  3735 CTAGAAATTCTGGCGGTTCAGTCAGTCGACTCTTGATACAAAAACCTACTGTGCCTAACGAGT 3794
         ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db  3547 CTAGAAATTCTGGCGGTTCAGTCAGTCGGCCAGATTAGGGTTAGGCCGTTGGTCGTGTCTGGAT 3606

Qy  3795 ACCATCCCCAGTACACAGCACATGCCACTCTTGATACAAAACTACTGTGCCTAACGAGT 3854
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3607 ACCATCCCCAGTACACAGCACATGCCACTCTTGATACAAAAACTACTGTGCCTAACGAGT 3666

Qy  3855 ATTCAGTTGCAAATTTAATTGCCCCCCACTGGCAGCGGCGCAAGTCAACCAAATTACCACTTT 3914
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3667 ATTCAGTTGCAAATTTAATTGCCCCCCACTGGCAGCGGCGCAAGTCAACCAAATTACCACTTT 3726

Qy  3915 CTTACATGCAGGAGAAGTATGAGTCTTGGTCCTAAATCCCAGTGTGGCTACAACAGCAT 3974
         |||||||||||||||||||||||:||||::|||||||||||||||||||||||||||||
Db  3727 CTTACATGCAGGRGAAGYATGAGYATGGGTGGTCCTAAATCCCAGTGTGGCTACAACAGCAT 3786

Qy  3975 CAATGCCAAAGTACATGCACGCGACGACGTCGGCCGCGTGAATCCAAATTGCTATTTTAATGCA 4034
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3787 CAATGCCAAAGTACATGCACGCGACGACGTACGCGCCGCGTGAATCCAAATTGCTATTTTAATGCA 3846

Qy  4035 AATGTACCAACACACAGAGGGGCTTCACTTACGTACGACACATATGCACATGTGACCGGAG 4094
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3847 AATGTACCAACACACAGAGGGGCTTCACTTACGTACGACACATATGCACATGTGACCGGAC 3906

Qy  4095 CATGTTCCCGGAACTATGATGTAATCATTGTGACGAATGCCATGCTACCGGTGCCAACCA 4154
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3907 GATGTTCCCGGAACTATGATGTAATCATTGTGACGAATGCCATGCTACCGGTGCGAACCA 3966

Qy  4155 CCGTGTTGGGCATTGGAAAGGTCCTAACCGAAGTCCATCCAAAAATGTAGGCTAGTGG 4214
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3967 CCGTGTTGGGCATTGGAAAGGTCCTAACCGAAGTCCATCCAAAAATGTAGGCTAGTGG 4026

Qy  4215 TTCTTGCCACGGCTACCCCCCTGGAGTAATCCCTACACCACATGCCAACATAACTGAGA 4274
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4027 TTCTTGCCACGGCTACCCCCCCTGGAGTAATCCCTACACCACATGCCAACATAACTGAGA 4086

Qy  4275 TTCAATTAACCGATGAAGGCACTATCCCTTTCATGGAAAAAGATTAAGGAGGAAAATC 4334
         |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4087 TTCAATTAACYGATGAAGGCACTATCCCTTTCATGGAAAAAGATTAAGGAGGAAAATC 4146

Qy  4335 TGAAGAAGGGAGACACCTTATCTTTGAGGCTACCAAAAACACTGTGATGAGCTTGCTA 4394
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4147 TGAAGAAGGGAGACACCTTATCTTTGAGGCTACCAAAAACACTGTGATGAGCTTGCTA 4206

Qy  4395 ACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGGATGTGACATCTCAA 4454
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4207 ACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGGATGTGACATCTCAA 4266

Qy  4455 AAATGCCTGAGGGCGACTGTGTAGTAGTTGCCACTGCCTTGTGCTTGTGTACAGGGTACTG 4514
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4267 AAATGCCTGAGGGCGACTGTGTAGTAGTTGCCACTGCCTTGTGCTTGTGTACAGGGTACACTG 4326

Qy  4515 GTGACTTTGATTCCGTGTATGACTCGCAGCCTCATGGTAGAAGCCACATGCCATGTTGACC 4574
         ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db  4327 GTGACTTTGATTCCGTGTATGACTCGCAGCCTCATGGTRGAAGCCACATGCCATGTTGACC 4386

Qy  4575 TTGACCCTACTTTCACCATGGGTGTTCGTTCGTGTGCGGGGTTTCAGCAATAGTTAAAGGCC 4634
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4387 TTGACCCTACTTTCACCATGGGTGTTCGTTCGTGTGCGGGGTTTCAGCAATAGTTAAAGGCC 4446

Qy  4635 AGCGGTAGGGGGCCCACAGCCGTGGGAGAGCTGGCCATATACTACTATGCTAGACGGGAGTT 4694
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4447 AGCGGTAGGGGGCCCACAGCCGTGGGAGAGCTGGCCATATACTACTATGCTAGACGGGAGTT 4506

Qy  4695 GTACCCCTTCGGGTATGCTTCGGTAGGCCTTCGACGCAGCCAAGG 4754
         |||||||||||||||||||||||||||||||||||||||||||||
Db  4507 GTACCCCTTCGGGTATGCTTCGGTAGGCCTTCGACGCAGCCAAGG 4566
```

```
Qy  5835 CTTTGACAAACAGCAGGGCCAGATCACTGGCCCAACAGACTTCTTACTATGCTTGCTAGGA 5894
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5647 CTTTGACAAACAGCAGGGCCAGATCACTGGCCCAACAGACTTCTTACTATGCTTGCTAGGA 5706

Qy  5895 GCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCGCAGGAAGATACTGG 5954
         |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db  5707 GCAACACTGTATGTAATGAGTACTTTAATTGCTTCGTGCACTCGTGACATCCGCAGGAAGATACTGG 5766

Qy  5955 GCATTCTGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCATCCGTTGGCTCCACA 6014
         ||||||||||||||||||||| ||| |||||| |||||
Db  5767 GCATTCTGGAGGCATCTACCCCCTGGAGTRTCATATCAGCTTGCATCCGTTGGCTYCACA 5826

Qy  6015 CCCCGACGGAGGATGATTGCGGCCTCATTGCGGCCTCAGAGATTTGGCAGTATGTGT 6074
         ||||||||||||||||||||||||||||||||||| |||||||||||||||
Db  5827 CCCCGACGCGGAGGATGATTGCGGCCTCATTGCGGCCTCARAGATTGGCAGTATGTGT 5886

Qy  6075 GCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGCTGGAGTTCAGAGCATGGTAACA 6134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5887 GCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGCTGGAGTTCAGAGCATGGTAACA 5946

Qy  6135 TTCCTGTGTTCCTTTCTACAGCTGCCAGAGGGGTACAAGGGCCCCTGGGATGGGATCAG 6194
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5947 TTCCTGTGTTGCCTTTCTACAGCTGCCAGAGGGGTACAAGGGCCCCTGGGATGGGATCAG 6006

Qy  6195 GTATGCTTCCAAGCACGTGTCCATGCGGTGCTGAACCTCATCTTTTCTGTTGGAATGGTT 6254
         |||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db  6007 GTATGCTTCCAAGCACGCGTGTCCATGCGGTGCTGAACCTCATCTTTTCTGTTGGAATGGTT 6066

Qy  6255 TTGCAAAACTTTACAAAGGACAGACCCAGGAACTTGTTCAAATTACTGGAGAGGGGCTGTTCCAG 6314
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  6067 TTGCAAAACTTTACAAAGGACAGACCCAGGAACTTGTTCAAATTACTGGAGAGGGGCTGTTCCAG 6126

Qy  6315 TCAACGGTAGGCTGTGTGGGGTCGGGTTCAGTGTTATCTAGACCCAACTAGTCTTGTCTG 6374
         |||||||||||||||||||| |||||||||||||||
Db  6127 TTCAACGGCTAGGCTGTGTGTGGGGTCGGGTTCAGTGTTATCTAGACCCAACTAGTCTTGTCTG 6186

Qy  6375 TCAATTATGGCGTTAGGGACTAGTACTGTAAAATGAGAAAATGGGAGATCACATTTTTGTTA 6434
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6187 TCAATTATGGCGTTAGGGACTAGTACTGTAAAATGAGAAAATGGGAGATCACATTTTTTGTTA 6246

Qy  6435 CAGCAGTATGATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCCAACCTTGAGAGCTGCAG 6494
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  6247 CAGCAGTATGATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCCAACCTTGAGAGCTGCAG 6306

Qy  6495 TGGCCGTGGACGGGCTACAAGGTTCAGTGTTATCTAGGTGAGGCCAAAACTCCTTGGACGA 6554
         |||||||||||||||||||||| ||||||||||||||||||||||||||
Db  6307 TGGCCGTGGACGGGCTACAAGGGTTCAGTGTTATCTAGGTGAGGCCAAAACTCCTTGGACGA 6366

Qy  6555 CATCTGCTTGCTGTGTTAGGGACGGTCCTGACGCGGGTAAGGGTAAAACTGTTAAGCTTCAGCC 6614
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  6367 CATCTGCTTGCTGTGTTAGGGACGGTCCTGACGCGGGTAAGGGTAAAACTGTTAAGCTTCCGCG 6426

Qy  6615 TTGACGGTCACACACTGGTCGTGCCATGCCAATTGCTTAATTGCTTGATGCACTTGAGACAA 6674
         |||||||||||||||||||||||||| |||||||||||||||||||||
Db  6427 TTGACGGTCACACACTGGTCGTGCCATGCCAATTGCTTAATTGCTTGATGCACTTGAGGCAA 6486

Qy  6675 ATGACTGTAATTCCACAAAACACACTCCTAGTGATGAAGCCGCACTTGAAGTCCGGCGTCTGTTT 6734
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  6487 ATGACTGTAATTCCATAAAACATCACTCCTAGTGATGAAGCCGCACTTGAAGTCCGGCGTCTGTTT 6546

Qy  6735 TTGACGGTTCGGTTCGCTTACCTTGCCTCCCCGAGATCCGTCCGGACTTCAGCTGGGACA 6794
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  6547 TTGACGGTTCGGTTCGCTTACCTTGCCTCCCCGAGATCCGTCCGGACTTCAGCTGGGACA 6606

Qy  6795 CCACCAAATGCCAGCCCCCCTCC---ATCGAAGAGGTAGTTGCTACTTCAGTTCC 6851
         ||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db  6607 CCACCAAATGCCAGCCCCCCTCCCCGAGATCGAAGAGGTAGTTGCTGTAAGAAAGCGCCAGTTCC 6666

Qy  6852 GGGCAAGAACTTGGTTCGCTTACCTTGCCTCCCCGAGATCCGTCCCAGGAGTGTCAT 6911
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  6667 GGGCAAGAACTTGGTTCGTTCGCTTACCTTGCCTCCCCGAGATCCGTCCCAGGAGTGTCAT 6726

Qy  6912 GTCCTGAAAGCCTGCAACGAAGTGACCCGGTTAGAAGGTCCTTCAAACCTCCCCTTCAC 6971
         |||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy       |||||||||||||||||||||||||||||||||| :  ::|||||||||||||||||||||| 6786
         GTCCTGAAAGCCTGCAACGAAGTGACCCGGTTAGAAGGTCCTTCAAMCCTCCCCTTCTTCAC
Db  6727

Qy  6972 CACCTGTTCTACAGTTGGCCATGCCGATGCCCGTGTTGGGACGCGGGTGAGTGATTGATTACCCAGTT 7031
         |||||||||||||||| :: |||||||||||||| : :|||||||||||||||||
Db  6787 CACCTGTTCTRACAGTTGGCCATGCCGATGCCCGTGTTGGGACGCAGGTGAGTGATTGATTTACCCAGTT 6846

Qy  7032 TCACTGCAATTGGATGTGCAATGCGCGAAAACAGGACCGGAGGGCCCTGATGATTACGGTT 7091
         ||||||||||||||||||||||| :: |||||||||||||||
Db  6847 TCACTGCAATTGGATGTGCAATGCGCGAAAACARGYGGAGKCCCWMAKRATTTACCCAGTT 6906

Qy  7092 ACCCTCCAAAAAGGAGAGTCTCTGAAATGTCAGACGAAAGTTGGTCGACGAAAGTTGGTGCGTACAACCG 7151
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  6907 ACCCTCCAAAAAGGAGAGTCTCTGAAATGGTCAGACGAAAGTTGGTCAGACGAAAGTTGGTCAACGACTACAACCG 6966

Qy  7152 CTTCCAGCTACGTTACTGGCCCCCCGTACCCCGTAAGATACGGGGAAAAGGGATTCCACTCAGT 7211
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  6967 CTTCCAGCTACGTTACTGGCCCCCCGTACCCCGTAAGATACGGGCAAGGAATTCCACTCAAT 7026

Qy  7212 CAGCCCCCGCCAAACGGCCTACAAAAGAGTGGAAAAAGAGTGAGTTTCGTGCAGCA 7271
         |||| |||||||||||||||||||||||||||||||||||||||||||||||
Db  7027 CAGCCCACCGCCAAACGGCCTACAAAAAGAGTGGGAAAAGAGTGAGTTTCGTGCAGCA 7086

Qy  7272 TGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGTCTCTGTCTGCAA 7331
         |||||||||||||||||||||| |||||| |||||||||||||
Db  7087 TGAGCTACACTTGGACCGACGTGATTAGCTTCAAAAACTGCTTCTAAAGTCTCTGTCTGCAA 7146

Qy  7332 CTCGGGCCATCACTAGTTGGTTTCCTCAAACAAAGATCATTGGTATGTGAACTGTGACTGAGCCGC 7391
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  7147 CTCGGGCCATCACTAGTTGGTTTCCTCAAACAAAGATCATTGGTATGTGAACTGTGACTGAGCCGC 7206

Qy  7392 GGGATGCGGAGCTAGAAAACAAAAAGTCACTATTAATAGACAACCTCTGTTCCCCCAT 7451
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  7207 GGGATGCGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCTGTTCCCCCAT 7266

Qy  7452 CATACCACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGTGCGGTGTCATGT 7511
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  7267 CATACCACCAAGCAAGTGAGATTGGCTAAGGAAAAAAGCTTCAAAAGTTGTGCGGTGTCATGT 7326

Qy  7512 GGGACTATGATGAAGTAGCAGCTCACACGCCCCTCGCTCAAGTCTGCTCAGTCCCATCACTG 7571
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  7327 GGGACTATGATGAAGTAGCAGCTCACACGCCCCTCGCTCAAGTCTGCTCAGTCCCATCACTG 7386

Qy  7572 GCCTTCGGGGCACTGATGTTCGTTCGTTCTGGGAGCGGCCCAAGCGTGTTCTGCAGA 7631
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  7387 GCCTTCGGGGCACTGATGTTCGTTCGTTCTGGGAGCGGCCCAAGCGTGTTCTGCAGA 7446

Qy  7632 AGTGTGTCGAGGCAGGTGAGATACCGGTCATTATCGGCAAACTGTGCAAACTGTGTAGTTCCAAAGG 7691
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  7447 AGTGTGTCGAGGCAGGTGAGATACCGGTCATTATCGGCAAACTGTGCAAACTGTGTAGTTCCAAAGG 7506

Qy  7692 AGGAGGTCTTCGTGAAGACCCCCAGAAAACCAACAAAGAAAACCCCCAGAAAACCAACAAAGAAAACCCC 7751
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  7507 AGGAGGTCTTCGTGAAGACCCCCAGAAAACCAACAAAGAAAACCCCCAGAAAACCAACAAAGAAAACCCC 7566

Qy  7752 ACCCCCACCTTGAAATGAGAAATGAGATGTTGAGAAGATGTGTTGAGAAGATGTGTACGGTACTACGG 7811
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  7567 ACCCCCACCTTGAAATGAGAAATGAGATGTTGAGAAGATGTGTTGAGAAGATGTGTACGGTACTACGG 7626

Qy  7812 TAGTTAAAGCTGTCATGGAGATCGTCATGGAGATCGCTCCACGTACCGGTTTGTAGATCCGTCAAGC 7871
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  7627 TAGTTAAAGCTGTCATGGAGATCGTCATGGAGATCGCTCCACGTACCGGTTTGTAGATCCGTCAAGC 7686

Qy  7872 GTCTGTTGTCGATGTGTCACCCGATCGGAGCGCCACATGCGAATCGGAGCGCCACATGCGAATACAGTGTTTTG 7931
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  7687 GTCGTGTTGTCGATGTGTCACCCGATCGGAGCGCCACATGCGAATCGGAGCGCCACATGCGAATACAGTGTTTTG 7746

Qy  7932 ACAGTACCATCACACCCGAGGATATCATGGTGAGACATCTACTTCAGCGACTAAAC 7991
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  7747 ACAGTACCATCACACCCGAGGATATCATGGTGAGACATCTACTTCAGCGACTAAAC 7806

Qy  7992 TCAGTGACCAACACCGAGTGCAACGAAGTGACCCGGTTAGAAGGTCTTATACGGTTATGGAGGAC 8051
         |||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 11
AAT00040
ID   AAT00040 standard; DNA; 8912 BP.
XX
AC   AAT00040;
XX
DT   27-AUG-2003 (revised)
DT   02-JUL-1996 (first entry)
XX
DE   Hepatitis GB virus (HGBV-B) DNA.
XX
KW   Hepatitis GB virus; HGBV; diagnosis; treatment; vaccine; reagents; non-A;
KW   non-B; non-C; non-D; non-E; HGBV-B; tamarin; infected plasma;
KW   lambda phage; cDNA library; ss.
XX
OS   Hepatitis G virus.
XX
FH   Key   Location/Qualifiers
FT   mat_peptide   4524..4751
FT   /*tag= a
FT   /note= "clone 48-1A1.1 protein prod."
FT   mat_peptide   6450..6731
FT   /*tag= b
FT   /transl_except= pos:6474..6476, aa:Ala
FT   /transl_except= pos:6630..6632, aa:Deleted
FT   /note= "clone 70-3A1.37 protein prod."
FT   misc_difference 6772
FT   /*tag= d
FT   /note= "given as j in specification, may be Stanford
FT   Ambiguity Code for C or A"
FT   mat_peptide   6834..7457
FT   /*tag= c
FT   /note= "clone 4-3B1.1 protein prod."
XX
PN   WO9521922-A2.
XX
PD   17-AUG-1995.
XX
PF   14-FEB-1995;   95WO-US002118.
XX
PR   14-FEB-1994;   94US-00196030.
PR   13-MAY-1994;   94US-00242654.
PR   29-JUL-1994;   94US-00283314.
PR   23-NOV-1994;   94US-00344185.
PR   23-NOV-1994;   94US-00344190.
PR   27-JAN-1995;   95US-00344557.
XX
PA   (ABBO) ABBOTT LAB.
XX
PI   Simons JN, Pilot-Matias TJ, Dawson GJ, Schlauder GG, Desai SM;
PI   Leary TP, Muerhoff AS, Erker JC, Buijk SL, Mushahwar IK;
XX
DR   WPI; 1995-293123/38.
DR   P-PSDB; AAR81402, AAR81403, AAR81404.
XX
PT   Non-A, non-B, non-C, non-D, non-E Hepatitis virus reagents - useful for
PT   diagnosis and therapy of hepatitis GB virus.
XX
PS   Example 4; Page 179-184; 661pp; English.
XX
CC   Double stranded hepatitis GB virus (HGBV) DNA obtd. from HGBV infected
CC   tamarin plasma, using standard procedures, was used to prepare a lambda
CC   phage HGBV cDNA library. cDNA clones (which encode the proteins AAR81402-
CC   04) rescued from the lambda phage, were found to have nearly 100%
CC   sequence homology with the HGBV-B DNA sequence AAT00040. Reagents which
CC   comprise the HGBV DNA, or its protein prods. can be used for the
CC   diagnosis, therapy or in a vaccine to prevent HGBV infection. (Updated on
CC   27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 8912 BP; 2019 A; 2169 C; 2265 G; 2400 T; 0 U; 59 Other;

Db   7807 TCAGTGACCAACACCGAGCTGGCATTCACACCATTGCCGAGGCAGTATCACGCTGGAGGAC 7866
Qy   8052 CGATGATCGCTTATGATGGCCCAGAGATCGGATATCGTGAGGTCTTCCGGCGTCT 8111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7867 CGATGATCGCTTATGATGGCCCAGAGATCGGATATCGTGAGGTGTAGGTCTTCCGGCGTCT 7926

Qy   8112 ATACTACCTCAAGTTCCAACAGTTTGACCTGCTGGCTGAGGTAAATGCTGCAGCCGAAC 8171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7927 ATACTACCTCAAGTTCCAACAGTTTGACCTGCTGGCTGAAGTGCTGCAGCCGAAC 7986

Qy   8172 AGGCTGGCATGAAGAACCCTCGCTTCCTTATTTGCGGCGATGATTGCACCGTAATTTGGA 8231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7987 AGGCTGGCATGAAGAACCCTCGCTTCCTTATTTGCGGCGATGATTGCACCGTAATTTGGA 8046

Qy   8232 AGAGCGCCGGAGCGAGATGCAGACAAAACAAGCAGCAATGCGTGTCTTGCTAGCTGGATGAAGG 8291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8047 AGAGCGCCGGACGCGGACGATGCAGACAAACAAGCAATGCGTGTCTTTGCTAGCTGGATGAAGG 8106

Qy   8292 TGATGGGTGCACCACAAGATTGTGTCGCCTCAAACCCAAATACAGTTTGGAAGAATTAACAT 8351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8107 TGATGGGTGCACCACAAGATTGTGTCGCCTCAAACCCAAATACAGTTTGGAAGAATTAACAT 8166

Qy   8352 CATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTACTACTTTCTTA 8411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8167 CATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTACTACTTTCTTA 8226

Qy   8412 CAAGAGATCCTGTATCCCCCTTGGCAGGTGCTCGCTGCCGAGGGTCTCGGATACAACCCCA 8471
              |||||||||||||||||||||||||||||||::|||||||||||||||||||||
Db   8227 CAAGAGATCCTGTATCCCCCTTGGCAGGTGCTCGCTGCCGAGGGTCTCGGATACAACCCCA 8286

Qy   8472 GTGCTGCCTGGATTGGGTATCTAATACATCACTACCCCATGTTGTGGGTTAGCCGTGTGT 8531
              |:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8287 GKGCKGCGTGGATTGGGTATCTAATACATCACTACCCCATGTTGTGGGTTAGCCGTGTGT 8346

Qy   8532 TGGCTGTGTCCATTTCATGGAGGACAGATGCTCTTTGAGGACAAACTTCCCGAGACTGTGACCT 8591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8347 TGGCTGTGTCCATTTCATGGAGGACAGATGCTCTTTGAGGACAAACTTCCCGAGACTGTGACCT 8406

Qy   8592 TTGACTGGTATGGGAAAAAAATTATACGGCCTGTGAGAGATCTCGCCCAGCCATCATTGCTG 8651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8407 TTGACTGGTATGGGAAAAAAATTATACGGCCTGTGCCTGAGAGATCTCGCCCAGCCATCATTGCTG 8466

Qy   8652 GTGTGCACGGTATTGAGGCTTTCTCGGTGTCGCGCTACACCACGCTGAGATCCTCAGAG 8711
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8467 GTGTGCACGGTATTGAGGCTTTCTCGGTGTCGCGCTACACCACGCTGAGATCCTCAGAG 8526

Qy   8712 TTTCCCAATCACTAACAGACATGACCATGCCCCCCCTGCGAGCCTGGCGAAAGAAAGCCA 8771
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8527 TTTCCCAATCACTAACAGACATGACCATGCCCCCCCTGCGAGCCTGGCGAAAGAAAGCCA 8586

Qy   8772 GGGCGGTGTCCTCGCCCAGCGCCAAGAGGGCCGGAGCACACCGCAAAATTGGCTGCGCTCC 8831
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   8587 GGGCGGTGTCCTCGCCCAGCGCCAAGAGGGCCGGAGCACACG-AAAATTGGCTGCGCTTCC 8645

Qy   8832 TTCTCTGCCATGCTACATCTAGACCTCTACACCACACGAGCGGTGCCTCCGT 8891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8646 TTCTCTGCCATGCTACATCTAGACCTCTACACCACACGAGCGGTGCCTCCGT 8705

Qy   8892 ACACCACTTTCAATTATTGTGATGTTTACTCCCCGGAGGGGATGTGTTTATTACACCAC 8951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8706 ACACCACTTTCAATTATTGTGATGTTTACT-CCCSGAGGRGGGATGTGTTTATTACACCAC 8764

Qy   8952 AGAGAAGATTGCAGAAGTTCCTTGAAGTAATTGGCTGTCATTGTTTTTGCCCTAGGGC 9011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8765 AGAGAAGATTGCAGAAGTTCTTGAAGTAATTGGCTGTCATTGTTTTGTGCCCTAGGGC 8824

Qy   9012 TCATTGCTGTTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAACTAACAGTTTT 9071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8825 TCATTGCTGTTGGACTAGCCATCAGCTGAACCCCCAAATTCAAAATTAACTAACAGTTTT 8884

Qy   9072 TTTTTTTTTTTTTTTTT 9090
              |||||||||||||||||
Db   8885 TTTTTTTTTTTTTTTTT 8903

Mon Oct 31 10:27:43 2005   Page 49

us-10-009-002-1.rng

```
Query Match      92.5%;  Score 8691.4;  DB 2;  Length 8912;
Best Local Similarity 98.7%;  Pred. No. 0;
Matches 8781;  Conservative 54;  Mismatches 54;  Indels 10;  Gaps 8;

Qy    196  TGGGTTCGGTGGTGGTGGTGTGGGGCTTTAGGCAGCCTCACGCCCACCACCTCCCAGATAGAGC  255
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11  TGGGTTCGGTGGTGGTGGTGTGGGGCTTTAGGCAGCCTCACGCCCACCACCTCCCAGATAGAGC   70

Qy    256  GGCGGCACTGTAGGGAGGAAGACCGGGACCCGGTCACTACCAAGGACGCAGACCTCTTTTTGA  315
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     71  GGCGGCACTGTAGGGAGGAAGACCGGGACCCGGTCACTACCAAGGACGCAGACCTCTTTTTGA  130

Qy    316  GTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTTGGGATGGTTGGGGT  375
           |||||||||||||||||||||||||||||||||||||: |||||||||||||||||||
Db    131  GTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACCTAYATGTTGGGATGGTTGGGGT  190

Qy    376  TAGCCATCCATACCGTACTGGCCTGATAGGGGATCTGGAGGGGATCTCGTAG  435
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    191  TAGCCATCCATACCGTACTGGCCTGATAGGGGATCTGGAGGGGATCTCGTAG  250

Qy    436  ACCGTAGCACATGCCTGTATTTCTACTCAAACAAGTCCTGTACCTGCGCCCAGAACGCG  495
           |||||||||||||||||||||||||: ||||||||: ||||||||: |||||||||||
Db    251  ACCGTAGCACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACCTGCGCRCCCAGAACGCG  310

Qy    496  CAAGAACAAGCAGCAGCCAGGCTTCATATCCTGTCTGTCCATTAAAACATCTGTTGAAAGGGG  555
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    311  CAAGAACAAGCAGCAGCCAGGCTTCATATCCTGTCTGTCCATTAAAACATCTGTTGAAAGGGG  370

Qy    556  ACAACGAGCAAAGTCAGGCGATGCTCGGCCCTGTAATTACAAAAAATTGCTGG  615
           |||||||: |: |||||||||||||||||||||||||||||||||||||||||
Db    371  ACAACGAGCAARGCGCAAAGTCCAGGCGATGCTCGGCCCTCGTAATTACAAAAAATTGCTGG  430

Qy    616  TATTCCATGATGGCTTGCAGACATTGGCTTCAGGCGATCCAGCTCAGGGGACG  675
           ||||||||||||||||||||||||||||||||: |||||||: |||||||||||
Db    431  TATTCCATGATGGCTTGCAGACATTGGCTTCAGGCGATCCAGCTCAGGGGACG  490

Qy    676  CCAAGACCCCTCGCCCATAAGTCTCGCAATCTTGGAATCCTTCTGGATTACCCTTTGGGGTG  735
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    491  CCAAGACCCCTCGCCCATAAGTCTCGCAATCTTGGAATCCTTCTGGATTACCCTTTGGGGTG  550

Qy    736  GATTGGTGATGTTACAACTCACACACCTCACACACCGTTGCTGCGAGAGGCCCGCTGGTGGCGGTCGT  795
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    551  GATTGGTGATGTTACAACTCACACACCTCACACACCGTTGCTGCGAGAGGCCCGCTGGTGGCGGTCGT  610

Qy    796  TCGGTGTCCACCTTTTGTGTATGTCTGCTATCTTGGGTATCTCCTGTATCTTTGGCCTGTGTAGTGGGC  855
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    611  TCGGTGTCCACCTTTTGTGTATGTCTGCTATCTTGGGTATCTCCTGTATYTTTGGCCTGTGTAGTGGGC  670

Qy    856  GCGGGTCACTGACCCAGACACAAATACCACAAATCCTGACCAATTGCTGCCAGCGTAATCA  915
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    671  GCGGGTCACTGACCCAGACACAAATACCACAAATCCTGACCAATTGCTGCCAGCGTAATCA  730

Qy    916  GGTTATCTATTGTTCTCCTTCACACCTGCCTACGAGCCTGGTCGTGTGATCGTGTCGCGGA  975
           ||||: |||||: |||||||||||||||||||||||||||||||||: ||: ||||||||
Db    731  GGTTATCTAYTGTYGTTCTCCTTCACACCTGCCTACGAGCCTGGTCGTGATCGTGTCGYGGA  790

Qy    976  CGAGTGCTGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAATTGGACTGGCACGGA  1035
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    791  CGAGTGCTGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAATTGGACTGGCACGGA  850

Qy   1036  CTCCTTCTTGGCTGACCACCATTGATTTTGTTATGGCGGCCTCTTGTGACCTGTGACGCCCT  1095
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    851  CTCCTTCTTGGCTGACCACCATTGATTTTGTTATGGCGGCCTCTTGTGACCTGTGACGCCCT  910

Qy   1096  TGACAATTGGTGAGTTGTGTGGTGTGTGTTGTGTGTAGTTAGTTCGGTGACTCGGTTCAGGCACTG  1155
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    911  TGACAATTGGTGAGTTGTGTGGTGTGTGTTGTGTGTAGTTAGTTCGGTGACTCGGTTCAGGCACTG  970

Qy   1156  GCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTGTACCTGGAAT  1216
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    971  GCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTGTACCTGGAAT  1030
```

```
Db   1031  GCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTGTACCTGGAAKTGCCTACTGGAAT  1090
           ||||||||||||||||||||||||||||||||||||||||||: |||||||||||||||
Qy   1276  AGATCCTCGGGTTCCTCAGGGTTTATCGGTGATGGCCGGCAAGGTCAGAGGTCGTCATCTT  1335

Qy   1276  AGATCCTCGGGTTCCTCAGGGTTTATCGGTGATGGCCGGCAAGGTCAGAGGTCGTCATCTT  1335
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1091  AGATCCTCGGGTTTATCGGGTTTATCGGTGATGGCCGGCAAGGTCAGAGGTCGTCATCTT  1150

Qy   1336  CTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTATGTTTAGCAGTGTACA  1395
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1151  CTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTATGTTTAGCAGTGTACA  1210

Qy   1396  CTACCTGGCGGTTGGCGCTCTGATCTACTATCGCTCTCGGGGCAAGTGGTATCAGTTGCT  1455
           ||||||||||||||||||||||||||||||||||: ||||: ||||||||||||||||||
Db   1211  CTACCTGGCGGTTGGCGCTCTGATCTAYGCCTCTAYGCCTCTCGGGGCAAGTGGTATCAGTTGCT  1270

Qy   1456  CCTAGCGCTTATGCTTTACATAGAAGCGACCTCTGGAAACCCATCAGGGTGCCCACTGG  1515
           |||||||||||||||||||||||||||||||||||||: |||||: ||||||||||||
Db   1271  CCTAGCGCTTAYGCTTTACATAGAAGCGACCTCTGGAAACCCYATCAGGGTGCCCACTGG  1330

Qy   1516  ATGCTCAATAGCTGAGTTTGCTCGCCTTGATGATACCCATGTCCTTGCCCACTCTTATTT  1575
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1331  ATGCTCAATAGCTGAGTTTGCTCGCCTTGATGATACCCATGTCCTTGCCCACTCTTATTT  1390

Qy   1576  GAGTGAGAATGTGTCAGAAGTCATTTGTTACAGTCCAAAGTGACCAGGCCTGATCCACTCT  1635
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1391  GAGTGAGAATGTGTCAGAAGTCATTTGTTACAGTCCAAAGTGACCAGGCCTGTCCACTCT  1450

Qy   1636  AGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCTGGTGCCGAGGGGATGTAT  1695
           ||||||||: |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1451  AGAGTATAABAACTCCATATCTTGGTACCCCTATACAATCCCGGTGCGGAGGGGATGTAT  1510

Qy   1696  GGTAAATTCAAAAATAACACATGGGGTTGCTGCCGATTCCGATGTGCCATCGTACTG  1755
           |||||||||||||||||||||||||||||: ||: |||||||||||||||||||||
Db   1511  GGTAAATTCAAAAATAACACATGGGGTTGCTGCCG-WTCGCAATGTGCCATCGTACTG  1569

Qy   1756  CACTATGGCACTGATGCAGTGTGGAACGACACTCGCAACACTTACGAGCATGCGGTGT  1815
           ||||||||||||||||||||||||||: |||||||||||||||||||||||||||||
Db   1570  CACTATGGCACTGATGCAGTGTGGAASSACAGTCGCAACACTTACGAGCATGCGGTGT  1629

Qy   1816  AACACCATGGCTAACAACCGCATGGCACAACCGCTCAGCCCTGAAATTGGCTATATTACA  1875
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1630  AACACCATGGCTAACAACCGCATGGCACAACCGCTCAGCCCTGAAATTGGCTATATTACA  1689

Qy   1876  ATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATGTCAGGCCATTGTATTT  1935
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1690  ATACCCTGGGTCTAAAGAAATGTTAAACCTCATATATTGTCAGGCCATTGTATTT  1749

Qy   1936  TGAGGGATCAGATACCCCTAGTAGTTTACTTTATGACCCGTGAATTCAGGCTCCTACC  1995
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1750  TGAGGGATCAGATACCCCTAGTAGTTTACTTTATGACCCGTGAATTCCACTCCTACC  1809

Qy   1996  ACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCACCTGTTCTTGGTTACA  2055
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1810  ACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCACCTGTTCTTGGTTACA  1869

Qy   2056  GGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAAGGATTGATCACCAAAGA  2115
           ||||||||: |: ||||||||||||||||||||||||||||||||||||||||||||||
Db   1870  GGTTCCGCAA-GGTTTTACAGTGATGTGAAAGACCTAGCCACAAGGATTGATCACCAAAGA  1928

Qy   2116  CAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGCTTTGTCTCTTACGGG  2175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1929  CAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGCTTTGTCTCTTACGGG  1988

Qy   2176  AGTTACCACCAAGGCCGTGGTTGCTAATTCTGTTGGGGTTGTGTGGCAGCAAGTATCTTAT  2235
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1989  AGTTACCACCAAGGCCGTGGTTGCTAATTCTGTTGGGGTTGTGTGGCAGCAAGTATCTTAT  2048

Qy   2236  TTTAGCCTTACCTGTCCCTTTGTTTTGGGCGGGCGCTTCTCGGTTACCCCTTTTGCG  2295
           ||||||||||||||||||||: ||: ||||||||||||||||||||||||||: |||
Db   2049  TTTAGCCTTACCTGTCCCTTYGTTTTGGGGCGGGCGCTTCTCGGTTACMCTTTGCG  2108

Qy   2296  TCCTGCTGCCCATCCCAGTCGTATCTTCCAAGCTGGGAAGTGCCCACTGGAAT  2355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

us-10-009-002-1.rng

```
Qy  3435 ATGGCAGCAAGGGGCGCCGGTTGGCTCATCCCACCAATACACCCAATAACCGTTG 3494
         ||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db  3247 ATGGCAGCAACGGGGCGCCGGTTGGCTCATCCCACCAGGCTCCATACACCCAATAACCGTTG 3306

Qy  3495 ACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGGGTCCCTTACTCGGT 3554
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3307 ACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGGGTCCCTTACTCGGT 3366

Qy  3555 GCTCTTGCGGGAGACCAAGGGGTATCTGGTAACACGACTGGGGTCATTGGGTCAGGGTCA 3614
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3367 GCTCTTGCGGGAGACCAAGGGGTATCTGGTAACACGACTGGGGTCATTGGGTCAGGGTCA 3426

Qy  3615 ACAAATCCGATGACCCTTATTGGTGTGTGTGCGGGGCCCTTCCCATGGCTGTTGCCAAGG 3674
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3427 ACAAATCCGATGACCCTTATTGGTGTGTGTGCGGGGCCCTTCCCATGGCTGTTGCCAAGG 3486

Qy  3675 GTTCTTCAGGTGCCCGATTCTGCGTCCTCCGGCATGTAATTGGGATGTTCACCGCTG 3734
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3487 GTTCTTCAGGTGCCCGATTCTGCGTCCTCCGGCATGTAATTGGGATGTTCACCGCTG 3546

Qy  3735 CTAGAAATTCTGGCGGTTCAGTCAGTCAGATTAGGGTTAGGCGCGGTGTGTGTGCTGAT 3794
         |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db  3547 CTAGAAATTCTGGCGGTTCAGTCAGTCCAGATTAGGGTTAGGCGCGGTGTGTGTGCTGAT 3606

Qy  3795 ACCATCCCCAGTACACAGCACATGCCACTCTGATACAAAACCTACTGTGCCTAACGAGT 3854
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3607 ACCATCCCCAGTACACAGCACATGCCACTCTGATACAAAACCTACTGTGCCTAACGAGT 3666

Qy  3855 ATTCAGTGCAAATTTTAATTGCCCCACTGGCAGCGGCAAGTCAACCAAATTACCACTTT 3914
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3667 ATTCAGTGCAAATTTTAATTGCCCCACTGGCAGCGGCAAGTCAACCAAATTACCACTTT 3726

Qy  3915 CTTACATGCAGGAGAAGTATGAGGTCTTGGTCTCCTAAATCCCAGTGTCGGTACAACAGCAT 3974
         ||||||||||||||| :||||||||||||||||||||||||||||||||||||||||
Db  3727 CTTACATGCAGGRGAAGYATGAGGTCTTGGTCTCCTAAATCCCAGTGTCGGTACAACAGCAT 3786

Qy  3975 CAATGCCAAAGTACATGCACGCGACGTACGGACGTACGGGTGAATCCAAATTGCTATTTTAATGGCA 4034
         |||||||||||||||||::|||||||||||||||||||||||||||||||||||||
Db  3787 CAATGCCAAAGTACATGCACGCGACGTACGGACGTACGGGTGAATCCAAATTGCTATTTTAATGGCA 3846

Qy  4035 AATGTACCACACAGCAGGGCGGCTTCACTACGCACATATGGCATGTACCGGAG 4094
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3847 AATGTACCACACAGCAGGGCGGCTTCACTACGCACATATGGCATGTACCGGAC 3906

Qy  4095 CATGTTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGCTACCGATGCAACCA 4154
         |||||||||||||||||||| ::|||||||||||||||||||||||||||||||||
Db  3907 GATGTTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGCTACCGATGCAACCA 3966

Qy  4155 CCGTGTTGGGCATTGGAAGGTCCTCAACGAAGCTCCATCCAAAAATGTTAGGCTAGTGG 4214
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3967 CCGTGTTGGGCATTGGAAGGTCCTCAACGAAGCTCCATCCAAAAATGTTAGGCTAGTGG 4026

Qy  4215 TTCTTGCCACGGCTACCCCCCCTGGAGTAATCCTACACCACATGCCAACATAACTGAGA 4274
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4027 TTCTTGCCACGGCTACCCCCCCTGGAGTAATCCTACACCACATGCCAACATAACTGAGA 4086

Qy  4275 TTCAATTAACCGATGAAGGCACTTATCTTTGAGCGTACCAAAAAACACTGTGATGAGCTTGCTA 4334
         ||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db  4087 TTCAATTAACYGATGAAGGCACTTATCTTTGAGCGTACCAAAAAAACACTGTGATGAGCTTGCTA 4146

Qy  4335 TGAAGAAAGGGAGACACTTATCTTACTAGGGGATGAGCACTGTGATGAGCATCTCAA 4394
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4147 TGAAGAAAGGGAGACACTTATCTTACTAGGGGATGAGCACTGTGATGAGCATCTCAA 4206

Qy  4395 ACGAGTTAGCTCGAAAGGGGAATAACAGCTTGTCTCTTACTAGGGGATGCCTTGTACAGGGTACACTG 4454
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4207 ACGAGTTAGCTCGAAAGGGGAATAACAGCTTGTCTCTTACTAGGGGATGCCTTGTACAGGGTACACTG 4266

Qy  4455 AAATCCCTGAGGGCGACTGTGTAGTTGCCACTGAGTCCTGTGTATACTGCTCACC 4514
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4267 AAATGCCTGAGGGCGACTGTGTAGTTGCCACTGAGTCCTGTGTATACTGCTCACC 4326
```

```
Db  2109 TCCTGTGCTCCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGTTTTGTCTAAAGCTCA 2168
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2356 AGTAGCTCCCTTTGCTTGATTTCTTCATCTGTGCTATCTCCGCTGCAGGCTACGTTA 2415
         ||||||:|||||||||||:|||||||||||||||||||||||||||||||||||||
Db  2169 AGTAGCTCMTTTTGTTTGATTTTCTTCATCTGTGCTATCTCCGCTGCAGGCTACGTTA 2228

Qy  2416 TGCTGCCCTTTTGTGCCATGGCTGCGCGGGCTTGCCCCTAACTTCTTGTTGC 2475
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2229 TGCTGCCCTTTTAGGGTTGTGCCATGGCTGCGCGGGCTTGCCCCTAACTTCTTGTTGC 2288

Qy  2476 AGCAGCTGCTGCCCAACCAGATTATGACTGGTTGGCGTGCGACTGCTAGTGGCAGGGGTTAGT 2535
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2289 AGCAGCTGCTGCCCAACCAGATTATGACTGGTTGGCGTGCGACTGCTAGTGGCAGGGGTTAGT 2348

Qy  2536 TTTGTGTGGCCGGCCGTAACCGTGGTCACCCGTACGCCGTGTCCTTGTGTAGTGTCCTTGGCCTCT 2595
         |||||||||||||||||||| |||||||::|||||||||||||||||||||||||||
Db  2349 TTTGTGTGGCCCGGCCGGCCGTGACCCGTGGTCA-CGCATAGCTCTGCTTGTGTAGTGTCCTTGGCCTCT 2407

Qy  2596 GGTAGCGCGTTT-TAACCCTCTTGCATTTGGTTACGCCGTCTTTGATACCGGAGA 2654
         ||||||||||||  ||||||||||||||||::|||||::|||||||||||||||
Db  2408 GGTAGCGCGTTTVTAACCCTCTTGCATTTSSTKACGCCCTGCTT-AGCTTTGACACCGGAGA 2466

Qy  2655 TAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCATGTCTCGTTTTGCT 2714
         ||||||||||||||||||||||||||||||||||||||::|||::|:|||||||||||
Db  2467 TAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCATGTCTCGTTTTGCT 2526

Qy  2715 TCTTTGCTCACTTGTTGTTACCTCGCTGCGTGCTTTAGTTAACTCCTATCTTTGGCAACGTTGGG 2774
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2527 TCTTTGCTCACTTGTTGTTACCTCGCTGCGTGCTTTAGTTAACTCCTATCTTTGGCAACGTTGGG 2586

Qy  2775 AGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTTCCTTGTGCTGGTTGTT 2834
         |||||||||||||||||||||||||||||:|||||||||||||||||:|||||||||
Db  2587 AGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTCCTTGTYGCTGGTTGTT 2646

Qy  2835 TCCCCGGTGCGACATATGAGCGCTGGTGACTTTCTGTGTGTCAGTAGCTCTTCTAT 2894
         |||||||||||||||||||:|:|||||||||||||||||||||||||||||||||
Db  2647 TCCCCGGTGCGACATATGAGCGGRGAAGYATGAGCWTTCTGTGTGTCAGTAGCTCTTCTAT 2706

Qy  2895 GTTTAACATCCAGTGCAGCATGTCTTGGGACTGACGTTAGGGCCCATAGAA 2954
         ||||||||||||||::::|||::|||||||||||||||||||||||||||
Db  2707 GTTTAACATCCAGTGCAGCAYMGTTCTTGGGACTGACGTTAGGGCCCATAGAA 2766

Qy  2955 TGTTGGTGCGTCTCGGAAAGTGTCATGCTTGGTATTCTCATTATGTTCTTAAGTTTTTCC 3014
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2767 TGTTGGTGCGTCTCGGAAAGTGTCATGCTTGGTATTCTCATTATGTTCTTAAGTTTTTCC 2826

Qy  3015 TCTTAGTGTTTGGTGAGAATGGTGTGTTTTCTATAGCACTTGCATGGTGATGTCTTGC 3074
         |||||||||||||||||||||||||||||::|||||||||||||||||||||||||
Db  2827 TCTTAGTGTTTGGTGAGAATGGTGTGTTTTCTAKAAGCACTTGCATGGTGATGTCTTGC 2886

Qy  3075 CTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCATTTTCCCCTTTGAAGGCAAGG 3134
         ||||||||||||||||||||||||||||||||||||::||||||||||||||||||
Db  2887 CTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCATTTTCCCCTTTGAAGGCAAGG 2946

Qy  3135 CAAGGGTTCTATAGGAATGAAGGAAGACGCTTGGCGTGTGGGGACACGGTTGATGGTTTGC 3194
         ||||||||||||||||||||||::|||||||||||||||::|||||||||||||||:
Db  2947 CAAGGGTTBGCGTTCTATAGGAATGAAGGAAGACGCTTGGSKKGTGGGGACACGGTTGATGGTTTGS 3006

Qy  3195 CCGTTGTTGCGCGTCTCGGCAGCTTTACCGGTCTCTCTGAACGTTGGCTATGCCGCGATGGGT 3254
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3007 SCGTTGTTGBGCGTCTCGGCAGCTTTACCGGTCTCTCTGAACGTTGGCTATGCCGCGATGGGT 3066

Qy  3255 GGGCCATTACCCGACCTTTACGGTCTATAGAACCGCGGGACTGGAACTGGACATCTTCAGGCA 3314
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3067 GGGCCATTACCCGACCTTTACGGTCTATAGAACTGGACTGGAACTGGACACGCGTGTCAGGCGA 3126

Qy  3315 TGGCAGTGGTCATGACTGGTATAGACCCCGAACTGGACTGGAACATCTTCAGATTAG 3374
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3127 TGGCAGTGGTCATGACTGGTATAGACCCCGAACTGGACTGGAACATCTTCAGATTAG 3186

Qy  3375 GATCTCTGCCACTAGCTGCCACATGGGATTTGTTTGTTGACAACGTGTTGTATACTGCTCACC 3434
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3187 GATCTCTGCCACTAGCTGCCACATGGGATTTGTTTGTTGAATGCGTGTTGAATACTGCTCACC 3246
```

```
Qy  4515  GTGACTTGATTCCGTGTATGACTGCAGCCTCATGTGAGAGGCACATGCCATGTGACC  4574
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4327  GTGACTTGATTCCGTGTATGACTGCAGCCTCATGTGAGAGGCACATGCCATGTGACC  4386

Qy  4575  TTGACCCTACTTCACCATGGGTGTTCGTGTGTGCGGGGTTCAGCAATAGTTAAAGGCC  4634
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4387  TTGACCCTACTTCACCATGGGTGTTCGTGTGTGCGGGGTTCAGCAATAGTTAAAGGCC  4446

Qy  4635  AGCGTAGGGGCGCACAGGCCGTGGGAGAGCCCGGACATATACTACTATGTAGAGGGAGTT  4694
          |||||||||| ||||||||| ||| ||||| ||||||||||||||||||||| | |||||
Db  4447  AGCGTAGGGGCCGCACAGGCCGGGGGAGAGCCGGACATATACTACTATGTAGACGGGAGTT  4506

Qy  4695  GTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAGAGCCTTGTGTCGACGCAGCCAAGG  4754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4507  GTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAGAGCCTTGTGTCGACGCAGCCAAGG  4566

Qy  4755  CATGGTATGGTTGTCATCAACAGAAGCTCAAACTATTCTGGACACTATCGCACCCCAAC  4814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4567  CATGGTATGGTTGTCATCAACAGAAGCTCAAACTATTCTGGACACTATCGCACCCCAAC  4626

Qy  4815  CTGGGGTTACCTGCGATAGGAGCAAATTGGACAGAGTGGGCTGATCTCTTTCTATGTCA  4874
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4627  CTGGGGTTACCTGCGATAGGAGCAAATTGGACAGAGTGGGCTGATCTCTTTCTATGTCA  4686

Qy  4875  ACCCCGAACCTTCATTTGTCAATACTGTCAAAAAGAACTGCTGACAATTATGTTTTGTTGA  4934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4687  ACCCCGAACCTTCATTTGTCAATACTGTCAAAAAGAACTGCTGACAATTATGTTTTGTTGA  4746

Qy  4935  CTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGTGCTCCCAATGACGCACCAC  4994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4747  CTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGTGCTCCCAATGACGCACCAC  4806

Qy  4995  GGTGGCAGGGAGCCCCGCTTGGGGAAAAACCTTGTGTGGGTTCGTGCGCTTGGACGGCG  5054
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db  4807  GGTGGCAGGGAGCCCCGCTTGGGGAAAAACCTTGTGTGGGTTCGTGCGCTTGGACGGCT  4866

Qy  5055  CTGACGCCTGTCCTGCGCCCAGCAGCGGAGGTGACCAGGATACCAAATGCTTCACTG  5114
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4867  GTGACGCCTGTCCTGCGCCCAGCAGCGGAGGTGACCAGGATACCAAATGCTTCACTG  4926

Qy  5115  AAGTCAATACTTCTGGGACAGCGCACTCGCCCACTGTCGTTGGCGTGTTGGCTTATC  5174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4927  AAGTCAATACTTCTGGGACAGCGCACTCGCCCACTGTCGTTGGCGTGTTGGCTTATC  4986

Qy  5175  TAGCCATTGACACTTTTGGGCGCCACTTGTGTCGCGGCCACTTGTGTCTATTACATCAGTCC  5234
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4987  TAGCCATTGACACTTTTGGGCGCCACTTGTGTCGCGGCCACTTGTGTCTATTACATCAGTCC  5046

Qy  5235  CTACCGGTGCTACTGCTCGCGCCCCAGTCGTTGACGAAGAAATCGTGGAGAGTGTGCAT  5294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5047  CTACCGGTGCTACTGCTCGCGCCCCAGTCGTTGACGAAGAAATCGTGGAGAGTGTGCAT  5106

Qy  5295  CATTCATTCCCTTGGAGGCCATGGTTGCTGCAATTGCTCGCAATTGCGAAGTTGACAAGCTGAAGAGTACAATCACCA  5354
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5107  CATTCATTCCCTTGGAGGCCATGGTTGCTGCAATTGCTCGCAATTGCGAAGTTGACAAGCTGAAGAGTACAATCACCA  5166

Qy  5355  CAACTAGTCCTTCACATTGAAAACCGCCCCTGAAAAACTTAACACCTTTCTTGGGCCTC  5414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5167  CAACTAGTCCTTCACATTGAAAACCGCCCCTGAAAAACTTAACACCTTTCTTGGGCCTC  5226

Qy  5415  ATGCAGCTACAAATCCTTGCTCATCAGTAGTATTGCTGTGTGTTTAGTCACTTACCTGACA  5474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5227  ATGCAGCTACAAATCCTTGCTCATCAGTAGTATTGCTGTGTGTTTAGTCACTTACCTGACA  5286

Qy  5475  ATCCCTTTGCATCATGCGTGTTTGCTTCATTGCGGGTATTGCCGGGTATTACACCCCACTACCTCACA  5534
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5287  ATCCCTTTGCATCATGCGTGTTTGCTTCATTGCGGGTATTGCCGGGTATTACACCCCACTACCTCACA  5346

Qy  5535  AGATCAAAATGTTCCTGTTCATTATTTGGAGGCGCAATTGCGTCCAAGCTTACAGACGCTA  5594
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5347  AGATCAAAATGTTCCTGTTCATTATTTGGAGGCGCAATTGCGTCCAAGCTTACAGACGCTA  5406

Qy  5595  GAGGGCACTGGCGTTCATGATGGGCCGGGGCTGCGGGAACAGCTCTTGGTACATGGACAT  5654
          ||||| |||||||||||||||||||||||||| | |||||||||||||||||||||||||
Db  5407  GAGRCGGCACTGGCGTTCATGATGGGCCGGGGCTGYGYGGAACAGCTCTTGGTACATGGACAT  5466

Qy  5655  CGGTGGGTTTGTCTTTGACATGTGATGGCGGCCTATGCTGGGCGGCCTCATCCCACTGCTTGCT  5714
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5467  CGGTGGGTTTGTCTTTGACATGTGATGGCGGCCTATGCTGGGCGGCCTCATCCCACTGCTTGCT  5526

Qy  5715  TGACATTTAAATGCTTAGTGGTGAGTGCCCACTACTGATCAGCTTGGTCTCGTGGTTTAGTCT  5774
          |||||||| ||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db  5527  TGACATTTAAATGCTTAGTGGTGAGTGCCYCACTACTGATCAGCTTGGTCTCGTGGTTTAGTCT  5586

Qy  5775  ACTCCGCGTTCAATCCGCCCGCCCGGGGCTCATCTGCGTCTTGTCGCGTCTTGTGCAACGTTTG  5834
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  5587  ACTCCCGCGTTCAATCCGCCCGCCCGGGGCTCATCTGCGTCTTGTCGCGTCTTGTGCAATGTTTG  5646

Qy  5835  CTTTGACAACAGCAGGGCCAGATCACTGGCCGCCCAACAGACTTCTTACTATGCTGCTAGGA  5894
          |||||||||||||||||||||||||||||| ||||||||||||||||||||| ||||||||
Db  5647  CTTTGACAACAGCAGGGCCAGATCACTGGCCCCCAACAGACTTCTTACTATGTGCTGCTAGGA  5706

Qy  5895  GCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCGCAGGAAGAGATACTGG  5954
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db  5707  GCAACACTGTATGTAARTGAGTACTTTATTGCCACTCGTGACATCCGCAGGAAGAGATACTGG  5766

Qy  5955  GCATTCTGGAGGCATCTACCCCCCTGGAGTGTCATATCAGCTTGGTCGTTGGCTCCACA  6014
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||| ||||
Db  5767  GCATTCTGGAGGCATCTACCCCCCTGGAGTGTRTCATATCAGCTTGGTCGTTGGCTYCACA  5826

Qy  6015  CCCCGACGCGGAGATGATTGCGGCCTCATTGCTTGGGGTTCCAGCTATGTGT  6074
          |||||||||||||||||||||||||||||||| ||||| | | | |||||||
Db  5827  CCCCGACGCGGAGATGATTGCGGCCTCATTGCTGGGGTCTARAGATTGGCAGTATGTGT  5886

Qy  6075  GCAATTCTTGTGATTTGCTTTAAATGTCCTCAGAGTGGAGTTCAGAGCATGGTTGTTAACA  6134
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5887  GCAATTCTTGTGATTTGCTTTAAATGTCCTCAGAGTGGAGTTCAGAGCATGGTTGTTAACA  5946

Qy  6135  TTCCTGGTGTCTTTCTACAGCTGCCAGATACCAAGGGGTACAAGGGCCCCTGGATTGGATCAG  6194
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5947  TTCCTGGTGTCCTTTCTACAGCTGCCAGATACCAAGGGGTACAAGGGCCCCTGGATTGGATCAG  6006

Qy  6195  GTATGCTCCAAGCACGCTGTCCAAGCAGCCGCACTGCGGTGCTGAACTCATCTTTCTGTGTTGAGAATGGTT  6254
          |||||||||||||||| ||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db  6007  GTATGCTCCAAGCACGGTGTCCAAGCGGCCGACGCGCTGCGGTGCTGAACTCATCTTTCTGTGTTGAGAGATGGTT  6066

Qy  6255  TTGCAAAAACTTACAAAGGACCAGAACTTGTTCAAATTACTACTGGAGAGGGGCTGTTCCAG  6314
          ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db  6067  TTGCAAAAACTTTACAAAGGACCAGAACTTGTTCAAATTACTACTGGAGAGGGGCTGTTCCAG  6126

Qy  6315  TCAACGGTAGGTGTGTGGGTCGGTCGAGCCCAACTGATTGACCGGACCCAACTGATTCGTCTTGTCG  6374
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||| ||||||||
Db  6127  TCAACGGTAGGTGTGTGGGTCGGTCGAGCCCAACTGATTGACCGGACCCAACTAGTCTTGTCG  6186

Qy  6375  TCAATTATGGCGTTAGGGACTACTGTAAAATGGAGCTGTAAAATGTGTTCACCCAGGTGAATCACATTTTTGTTA  6434
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6187  TCAATTATGGCGTTAGGGACTACTGTAAAATGGAGCTGTAAAATGTGTTCACCCAGGTGAATCACATTTTTGTTA  6246

Qy  6435  CAGCAGTATCCTCTCCAAATGTCTGTTCACCCAGGTGCCCCAACCTTGAGAGCTGCAG  6494
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6247  CAGCAGTATCCTCTCCAAATGTCTGTTCACCCAGGTGCCCCAACCTTGAGAGCTGCAG  6306

Qy  6495  TGGCCGGTGACGCGTACAGGTTCAGTGTTATCTGGACGGTGTTATCTGGTGAGCGGCCAAAACTCCTTGGACGA  6554
          |||||| |||||||||||||||||| ||||||||||||||||||||| ||||||||||||||||||||||||||
Db  6307  TGGCCGTGGACGCGTACAGGTTCAGYGTTATCTGGACGGTGTTATCTAGGTGAGCGGCCAAAACTCCTTGGACGA  6366

Qy  6555  CATCTGCTTGCTCCTGACGTAAGGCGTAAGGCTAAAACTGTTAAGCTTCCCTTCCGCG  6614
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db  6367  CATCTGCTTGCTCCTGACGTAAGGCGTAAGGCTAAAACTGTTAAGCTTCCCTTCCGGG  6426

Qy  6615  TTGACGGTCACACACCTGGTGTCGCGATGCAACTTAATTTGCGTGATGCACTTGAGACAA  6674
          |||||||  |||||||||||||||||||||||||||||||||||||||||||||||| ||
Db  6427  TTGACGGACACACACCTGGTGTCGCGATGCAACTTAATTTGCGTGATGCACTTGAGGCAA  6486

Qy  6675  ATGACTGTAATTCCACAAACAACACTCCTAGTGATGAGAAGCCCAGTGTCGCGCTCTTGTTT  6734
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  6487  ATGACTGTAATTCCATAAACAACACTCCTAGTGATGAAGCCGAGTGTCCGCTCTGTGTT  6546

Qy  6735  TCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTCAGCTGGCGTTGACA  6794
          ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||||
Db  6547  TCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTCAGCTGGCGTTGACA  6606

Qy  6795  CCACCAAACTGCCAGCCCCTCC---ATCGAAGAGAGTAGTGGTAAGAAAGCGCCAGTTCC  6851
          |||||||||||||||||||||||   |||||||||||||||||||||||||||||||||||
Db  6607  CCACCAAACTGCCAGCCCCTCCCCAGATCGAAGAGAGTAGTGGTAAGAAAGCGCCAGTTCC  6666

Qy  6852  GGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCCCGTCCCAGGAGTGTCAT  6911
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6667  GGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCCCGTCCCAGGAGTGTCAT  6726

Qy  6912  GTCCTGAAAGCTGCAACGAAGTGCTCCTTGAAGGTCCTTCAAACCTCCCTCCTTCAC  6971
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6727  GTCCTGAAAGCCTGCAACGAAGTGCACCGTTAGAAGGTCCTTCAANCCTCCCTTCTTCAC  6786

Qy  6972  CACCTGTTCTACAGTTGGCCATGCCGATGCCCGTTGGGAGCGGGTGAGTGTAACCCTT  7031
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6787  CACCTGTTCTRCAGTTGGCCATGCCGATGCCCGTTGGGAGCAGGTGAGTGTAACCCTT  6846

Qy  7032  TCACTGCAATTGGATGTGCAATGACCGAAACAGGCGGAGGCCCTGATTGATTTACCCAGTT  7091
          |||||||||||||||||||||||||||||||| |||   |||:::||||||||||||||||
Db  6847  TCACTGCAATTGGATGTGCAATGACCGAAACARGYGGAGKCCCNMAKRATTTACCCAGTT  6906

Qy  7092  ACCCTCCCAAAAAGGAGGTCTCTGAATGCTGAACGAGAAAGTTGGTCGACGGCGTACAACCG  7151
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6907  ACCCTCCCAAAAAGGAGGTCTCTGAATGCTGAACGAGAAAGTTGGTCGACGGCGTACAACCG  6966

Qy  7152  CTTCCAGTCACGTTACTGGCCCCCCGTACCCTAAGATACGGGAAAGGATTCCACTCAGT  7211
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  6967  CTTCCAGTCACGTTACTGGCCCCCCGTACCCTAAGATACGGGAAAGGATTCCACTCAAT  7026

Qy  7212  CAGCCCCCGCCAAACGGCCTACAAAAAGAAGTTGGGAAAGAAGTTGAGTTTCGTCGAGCA  7271
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7027  CAGCCACCGCCAAACGGCCTACAAAAAGAAGTTGGGAAAGAAGTGAGTTTCGTCGCACA  7086

Qy  7272  TGAGTACACCTGGACCGAGCGTGATTAGCTTCAAACTGCTTCAAAATGTTCGTCTGCAA  7331
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7087  TGAGTACACTTGGACCGAGCGTGATTAGCTTCAAACTGCTTCAAAATGTTCGTCTGCAA  7146

Qy  7332  CTCGGGGCCATCACTAGTTGGTTCCTCAAACAAAGATCATTGGTATGTGACTGAGCCGC  7391
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7147  CTCGGGGCACTCACTAGTTGGTTCCTCAAACAAAGATCATTGGTATGTGACTGAGCCGC  7206

Qy  7392  GGGATGCGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCTGTTCCCCCCAT  7451
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7207  GGGATGCGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCTGTTCCCCCAT  7266

Qy  7452  CATACCACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTCAAAAGTGTCATGT  7511
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7267  CATACCACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTCAAAAGTGTCATGT  7326

Qy  7512  GGGACTATGATGAAGCAGCTCACACGCTCTAAGTCTGCTAAGTCTGTGCTCCACATCACTG  7571
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7327  GGGACTATGATGAAGCAGCTCACACGCTCTAAGTCTGCTAAGTCTGTGCCCACATCACTG  7386

Qy  7572  GCCTTGCACGGACTGATGTTCGTTCGTTCTGGAGCAGCCCGCAAGGCGTTCTGGACTTGCAGA  7631
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7387  GCCTTGCACGGACTGATGTTCGTTCGTTCTGGAGCAGCCCGCAAGGCGTTCTGGACTTGCAGA  7446

Qy  7632  AGTGTGTCGAGGGACTGTGTTGAAGATGTGTTGAAGATGTACTACGGTCAGGTTGCTCCTGACG  7691
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7447  AGTGTGTCGAGGGACTGTGTTGAAGATGTGTTGAAGATGTACTACGGTCAGGTTGCTCCTGACG  7506

Qy  7692  AGGAGGTGTCTTCGTTGAAGACCCCCAGAACAACAAAGAAACCCCCCAAGGCTTATCTCGT  7751
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7507  AGGAGGTCTTCGTTGAAGACCCCCAGAACAACAAAGAAACCCCCCAAGGCTTATCTCGT  7566

Qy  7752  ACCCCACCTTGAAATGAGATGTGTTGAAGAGATGTACTACGGTCAGGTTGCTCCTGACG  7811
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7567  ACCCCACCTTGAAATGAGATGTGTTGAAGAGATGTACTACGGTCAGGTTGCTCCTGACG  7626
```

```
Qy  7812  TAGTTAAAGCTGTCATGGGAGATGCGTACGGGGTTTGTAGATCACGTACCCGTGTCAAGC  7871
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7627  TAGTTAAAGCTGTCATGGGAGATGCGTACGGGGTTTGTAGATCACGTACCCGTGTCAAGC  7686

Qy  7872  GTCTGTTGTCGATGTGGTCACCCGAGTCGAGCCACCACAGTGTGTTTTG  7931
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7687  GTCTGTTGTCGATGTGGTCACCCGAGTCGAGCCACCACAGTGTGTTTTG  7746

Qy  7932  ACAGTACCATCACACCCGAGATATCATCGTGGAGACACATCTACTCAGCAGTAAAC  7991
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7747  ACAGTACCATCACACCCGAGATATCATCGTGGAGACAGACATCTACTCAGCAGTAAAC  7806

Qy  7992  TCAGTGACCAAACACCGAGCTGGCATTCACACCATTGCGAGGCAGTTATACGCTGGAGGAC  8051
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7807  TCAGTGACCAAACACCGAGCTGGCATTCACACCATTGCGAGGCAGTATCACGCTGGAGGAC  7866

Qy  8052  CGATGATGCTTATGATGCCGGAGAGATCGGATATCGGATAGTCGTAGGTTCTTCCGGCGTCT  8111
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7867  CGATGATGCTTATGATGCCGGAGAGATCGGATATCGGATAGTCGTAGGTTCTTCCGGCGTCT  7926

Qy  8112  ATACTACCTCAAGTTCCAACAGTTTGACCTGCTCGGCTGGCTGAAGTAAATGCTGCAGCCGAAC  8171
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7927  ATACTACCTCAAGTTCCAACAGTTTGACCTGCTCGGCTGGCTGAAGTAAATGCTGCAGCCGAAC  7986

Qy  8172  AGGCTGGCATGAAGAACCCTCGCTTCCTTATTTGCGGCGGATGATTGCACCGTAATTTGGA  8231
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7987  AGGCTGGCATGAAGAACCCTCGCTTCCTTATTTGCGGCGGATGATTGCACCGTAATTTGGA  8046

Qy  8232  AGAGCGCCGGAGCGAGATGCAGATCAGACAAAACAAGCAATGCGGTCGTGTCTTTGCGTCTTGCGATGAAGG  8291
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8047  AGAGCGCCGGAGCGAGATGCAGATCAGACAAAACAAGCAATGCGGTCGTGTCTTTGCGTCTTGCGATGAAGG  8106

Qy  8292  TGATGGGTGCACCAAGATTGTGTGCCTCAACCCAAATACAGTTTGGAAGAATTAACAT  8351
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8107  TGATGGGTGCACCAAGATTGTGTGCCTCAACCCAAATACAGTTTGGAAGAATTAACAT  8166

Qy  8352  CATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTACTACTTTCTTA  8411
          ||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8167  CATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTACTACTTTCTTA  8226

Qy  8412  CAAGAGATCCTCGTATCCCCCTTGGCAGTGGCTCTCGGCGATACAACCCCA  8471
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8227  CAAGAGATCCTCGTATCCCCCTTGGCAGTGGCTCTCGGCGATACAACCCCA  8286

Qy  8472  GTGCTGCGTGGATTGGGTATCTAATCAATGATCACTACCCATGTTTGTGGGTTAGCCGTGT  8531
          ||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db  8287  GKGCKGCGTGGATTGGGTATCTAATCAATGATCACTACCCATGTTTGTGGGTTAGCCGTGT  8346

Qy  8532  TGGCTGTCATTTCATGGACAGATGCTCTTTGAGGACAAAACTTCCCGAGACTGTGACCT  8591
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8347  TGGCTGTCATTTCATGGACAGATGCTCTTTGAGGACAAAACTTCCCGAGACTGTGACCT  8406

Qy  8592  TTGACTGGTATGGGAAAAATTATACGGTGTCTTGCCTGTAAGCTCGTGCCCCAGCATCATTGCTG  8651
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8407  TTGACTGGTATGGGAAAAATTATACGGTGTCTTGCCTGTAAGCTCGTGCCCCAGCATCATTGCTG  8466

Qy  8652  GTGTGCACGGTATTGAGGCTTTCGGTGGTGTCGGTGTCGCCTACAACCCCAGATCCTCAGAG  8711
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8467  GTGTGCACGGTATTGAGGCTTTCGGTGGTGTCGGTGTCGCCTACAACCCCAGATCCTCAGAG  8526

Qy  8712  TTTCCCAATCACTAACAGACATGACCATGCCCCCCCCTGGCCAGAGAAAGCCA  8771
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8527  TTTCCCAATCACTAACAGACATGACCATGCCCCCCCCTGGCCAGAGAAAGCCA  8586

Qy  8772  GGGCGGGTCCTCGCCGCCAAGAGGCTGGCGGAGCACACGCGCAAAAATTGGCTCGCTTCC  8831
          |||||||||||||||||||||||||||||||||||||| -AAAATTGGCTCGCTTCC
Db  8587  GGGCGGGTCCTCGCCGCCAAGAGGCTGGCGGAGCACACG-AAAATTGGCTCGCTTCC  8645

Qy  8832  TTCTCTGGCATGCTACATCTGAACTCTACCGGTTGCTCCTGACG  8891
          ||||||||||||||||||||||||||||||||||||||||||||
Db  8646  TTCTCTGGCATGCTACATCTAGCCTCTACCGGTTGCTCCTGACG  8705
```

```
Qy  8892 ACACCACTTCAATTATTGTGATGTTTACTCCCCGAGGGGGATGTGTTATTACCAC 8951
         |||||||||||||||||||||||||||||  ||:||:|||||||||||||||||||
Db  8706 ACACCACTTCAATTATTGTGATGTTTACT-CCCSGAGRGGGATGTGTTTATTACACCAC 8764

Qy  8952 AGAGAAGATTGCAGAAGTTCCTTGTGAAGTATTGGCTGTCATTGTTTTTGCCCTAGGGC 9011
         |||||||||||||||||||  |||||||||||| |||||||||||||| ||||||||||
Db  8765 AGAGAAGATTGCAGAAGTTTCTTGTGAAGTATTTGGCTGTCATTGTTTGTGCCCTAGGGC 8824

Qy  9012 TCATTGCTGTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAACTAACAGTTTT 9071
         |||||||||||||| ||||||||||||||||| ||||||||||||| ||||| |||||
Db  8825 TCATTGCTGTGGACTAGCCATCAGCTGAACCCCCAAATTCAAAATTAATTAACACAGTTTT 8884

Qy  9072 TTTTTTTTTTTTTTTT 9090
         ||||||||||||||||
Db  8885 TTTTTTTTTTTTTTTTTT 8903
```

RESULT 12
ADA77751
ID   ADA77751 standard; RNA; 8069 BP.
XX
AC   ADA77751;
XX
DT   20-NOV-2003 (first entry)
XX
DE   Hepatitis GB virus B subgenomic neo-RepD replicon RNA sequence.
XX
KW   neo-RepD; GB virus-B; GBV-B; replicon; autonomous replication; NS3-NS5B;
KW   HCV; hepatitis C virus; antiviral; anti-HCV;
KW   neomycin phosphotransferase gene; neo; ss.
XX
OS   Hepatitis GB virus B.
XX

FH   Key   Location/Qualifiers
FH
FT   5'UTR   1..445
FT   /*tag= a
FT   CDS   446..1315
FT   /*tag= b
FT   /product= "Core-neo fusion protein"
FT   misc_signal   1324..1934
FT   /*tag= c
FT   /label= IRES
FT   /note= "IRES= Internal ribosome entry site of the
FT   encephalomyocarditis virus, drives translation of the GBV
FT   -B NS region"
FT   CDS   1935..7712
FT   /*tag= d
FT   /product= "GBV-B polyprotein NS3-NS5B"
FT   /note= "NS= non structural protein"
FT   CDS   5940..7712
FT   /*tag= e
FT   /partial
FT   /product= "Non structural protein 5B (NS5B)"
FT   /note= "NS5B is an RNA dependent RNA polymerase; Start
FT   codon is absent"
FT   3'UTR   7710..8069
FT   /*tag= f
XX
PN   WO200305944-A2.
XX
PD   24-JUL-2003.
XX
PF   13-JAN-2003; 2003WO-EP000281.
XX
PR   15-JAN-2002; 2002US-0348573P.
PR   06-JUN-2002; 2002US-0386655P.
XX
PA   (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
PI   De Tomassi A, Graziani R, Paonessa G, Traboni C;
XX
DR   WPI; 2003-598503/56.
XX

PT   New GB virus B (GBV-B) replicon for identifying compounds that inhibit
PT   GBV-B or hepatitis C virus, comprises a GBV-B 5' UTR, a selection or
PT   reporter sequence, an internal ribosome entry site, an NS3-NS5B sequence,
PT   and a GBV-B 3' UTR.
XX
PS   Claim 1; Fig 1; 81pp; English.
XX
CC   This invention relates to a novel GB virus-B (GBV-B) replicon and
CC   replicon enhanced cells. A GBV-B replicon is an RNA molecule able to
CC   autonomously replicate in a cultured cell to produce detectable levels of
CC   one or more GBV-B proteins. Specifically, it may comprise the GBV-B 5'
CC   UTR, GBV-B structural region, NS3-NS5B sequence, and GBV-B 3' UTR. Accordingly,
CC   they are useful in providing tools for studying GBV-B replication,
CC   polyprotein production and processing, identifying compounds that inhibit
CC   GBV-B, providing a surrogate model for identifying compounds that inhibit
CC   HCV, and providing a scaffold for producing GBV-B/HCV chimeric replicons.
CC   Due to the similarity between GBV-B and the hepatitis C virus (HCV),
CC   compounds that inhibit GBV-B may be useful antiviral agents, specifically
CC   anti-HCV agents. The GBV-B subgenomic replicon constructs termed GBV-B-
CC   neo-RepA (neo-RepA), neo-RepB, neo-RepC and neo-RepD were produced by
CC   replacing the regions coding for structural proteins and the NS2 protein
CC   with the sequences of neomycin phosphotransferase gene (neo) and
CC   encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) in
CC   the plasmid FL3/pACYC177. This polynucleotide sequence is the subgenomic
CC   neo-RepD replicon sequence of the invention.
XX
SQ   Sequence 8069 BP; 1863 A; 2075 C; 2136 G; 0 T; 1995 U; 0 Other;

Query Match   64.8%; Score 6088; DB 9; Length 8069;
Best Local Similarity   74.2%; Pred. No. 0;
Matches 4554; Conservative 1558; Mismatches 20; Indels 2; Gaps 1;

```
Qy  3266 GCACCTTTTACGCTGCAGTGTCTTCTGTGAACGTGGCACGCTGTCAGCGATGGCAGTGTC 3325
         |||||:::||||||||:||:||||:||:||:||||||||||:|:||::||||||||||||
Db  1938 GCACCUUUUACGCUGCAGUGUCUUCUGUGAACGUGGCACGCUGUCAGCGAUGGCAGUGUC 1997

Qy  3326 ATGACTGGTATAGACCCCGGAACTTGGACTGGAACTATCTTCAGATTAGGATCCTCGGCC 3385
         |:||||||||||||||||||||||:||||||||||:||:||||||:||||||:||:||||
Db  1998 AUGACUGGUAUAGACCCCGGAACUUGGACUGGAACUAUCUUCAGAUUAGGAUCCUCGGCC 2057

Qy  3386 ACTAGCTACATGGACGTGGATTGTTTGTGACAACGTGTTGTATACTGCTCACCATGGCAGCAAG 3445
         ||:|||:||||||||||||:||:||:||||||||||:||:|:||||:||:||||||||||||||
Db  2058 ACUAGCUACAUGGACGUGGAUUGUUUGUGACAACGUGUUGUAUACUGCUCACCAUGGCAGCAAG 2117

Qy  3446 GGGCGCCGGTTGGCTCATCCCACAGGCTCCTATACACCCAATAACCGTTGACCGCGGCTAAT 3505
         |||||||||:||||:||:|||||||||:||:|:||||||||:|||||:|||||||||||:||:
Db  2118 GGGCGCCGGUUGGCUCAUCCCACAGGCUCCUAUACACCCAAUAACCGUUGACCGCGGGCUAAU 2177

Qy  3506 GACCAGGACATCTATCAACCACCATGTGGAGCTGGGTCCCTTACTCGGGTGTCCTTGCGGG 3565
         |||||||||:|:|:||||||||||:|:||||||:|||:|||:|||:||||:|:||:|||||
Db  2178 GACCAGGACAUCUAUCAACCACCAUGUGGAGCUGGGUCCCUUACUCGGGUGUCCUUGCGGG 2237

Qy  3566 GAGACCAAGGGGTATCTGGTAACACGACTGGGGTCATTGGTTGAGGTCAACAAATCCGAT 3625
         |||||||||||||:|:|||:||||||||:||||||:||||:|||||||||||||:||||:
Db  2238 GAGACCAAGGGGUAUCUGGUAACACGACUGGGGUCAUUGGUUGAGGUCAACAAAUCCGAU 2297

Qy  3626 GACCCTTATTGGTGTGTGCGGGGCCCTTCCCATGGCTGTTGCCCAAGGGTTCTTCAGT 3685
         |||||:||:||||:|:||||||||||:||||||||:|||:|||||||||:||:||||:
Db  2298 GACCCUUAUUGGUGUGUGCGGGGCCCUUCCCAUGGCUGUUGCCCAAGGGUUCUUCAGU 2357

Qy  3686 GCCCGGATTCTGTTGCTCCTCCGGGCCATGTATTATTGGGATGTTCACCGCGTGCTAGAAATTCT 3745
         |||| :|||:||:||||||:||||||||||:||:|||||||:||:|||||||||:|||||||||:
Db  2358 GCCCCGAUUCUGUUGCUCCUCCGGGCCAUGUAUUAUUGGGAUGUUCACCGCGUGCUAGAAAUUCU 2417

Qy  3746 GGGGGTTCAGTCAGTCAGATTAGGGTTAGGCCGTTGGTGTGTCTGGATACCATCCCCAG 3805
         ||| |:|||||||| ||||||||||:|||||||:|||||:|:||||:|:|||:||||||
Db  2418 GGGCGGUUCAGUCAGCAGCAGAUUAGGGUUAGGCCGUUGGUGUGUCUGGAUUAACCAUCUCCCCAG 2477

Qy  3806 TACACAGCACATGCCACTCTTGATACAAAAACTACTGTGCCTAACGAGTATTCAGTGCAA 3865
         |||||||||||:||||:|:||||||||||||::|||:||||:|||||||:||:||||||
Db  2478 UACACAGCACAUGCCACUCUUGAUACAAAAACCUACUGUGCCUAACGAGUAUUCAGUGCA 2537
```

```
Db       ||||||||||||||||||||||||||||||||||||||||||||||| :||::||:::|:::||||||||||||||||||:
         7996 AGAGCGUAGUACCAAGGGCUGCACCCCGGUUUUUGUUCCAAGCGGAGGGCAACCCCCGCU 8055

Qy       9386 TGGAATTAAAACT 9399
         :||:::||:||||:
Db       8056 UGGAAUUAAAAACU 8069

RESULT 13
AAT00052
ID    AAT00052 standard; DNA; 4268 BP.
XX
AC    AAT00052;
XX
DT    27-AUG-2003 (revised)
DT    02-JUL-1996 (first entry)
XX
DE    Hepatitis GB virus (HGBV) clone.
XX
KW    Hepatitis GB virus; HGBV; diagnosis; treatment; vaccine; reagents; non-A;
KW    non-B; non-C; non-D; non-E; clone; tamarin; infected plasma;
KW    lambda phage; cDNA library; ss.
XX
OS    Hepatitis G virus.
XX
FH    Key             Location/Qualifiers
FT    mat_peptide     1..4266
FT                    /*tag= a
FT                    /label= AAR82066
FT    mat_peptide     complement(1..4266)
FT                    /*tag= d
FT                    /label= AAR82069
FT    mat_peptide     2..4267
FT                    /*tag= b
FT                    /label= AAR82067
FT    mat_peptide     complement(2..4267)
FT                    /*tag= e
FT                    /label= AAR82070
FT    mat_peptide     3..4268
FT                    /*tag= c
FT                    /label= AAR82068
FT    mat_peptide     complement(3..4268)
FT                    /*tag= f
FT                    /label= AAR82071
XX
PN    WO9521922-A2.
XX
PD    17-AUG-1995.
XX
PF    14-FEB-1995;    95WO-US002118.
XX
PR    14-FEB-1994;    94US-00196030.
PR    13-MAY-1994;    94US-00242654.
PR    29-JUL-1994;    94US-00283314.
PR    23-NOV-1994;    94US-00344185.
PR    23-NOV-1994;    94US-00344190.
PR    27-JAN-1995;    95US-00344557.
XX
PA    (ABBO ) ABBOTT LAB.
XX
PI    Simons JN, Pilot-Matias TJ, Dawson GJ, Schlauder GG, Desai SM;
PI    Leary TP, Muerhoff AS, Erker JC, Buijk SL, Mushahwar IK;
XX
DR    WPI; 1995-293123/38.
DR    P-PSDB; AAR82066, AAR82067, AAR82068, AAR82069, AAR82070, AAR82071.
XX
PT    Non-A, non-B, non-C, non-D, non-E Hepatitis virus reagents - useful for
PT    diagnosis and therapy of hepatitis GB virus.
XX
PS    Example 5; Page 232-234; 661pp; English.
XX
CC    Double stranded hepatitis GB virus (HGBV) DNA obtd. from HGBV infected
CC    tamarin plasma, using standard procedures, was used to prepare a lambda
```

```
Qy  8246 GATGCAGACAAACAAGCAATGCGTGTCTTTGCTCAGCTCGGATGAGGTGATGGGTGCACCA 8305
         ||||||||||||||||||||||:||| ::|:::|||::||:::::||||:|||||||||||
Db  6918 GAUGCAGACAAACAAGCAAUGCGUGUCUUUGCUCAGCUCGGAUGAGGUGAUGGGUGCACCA 6977

Qy  8306 CAAGATTGTGTGCCTCAACCCAAATACAGTTTGGAAGAATTAACATCATGCTTCATCAAAT 8365
         |||||||||||||:|||||::|:::||| ||::|||::::||| :|||||||:|||:||||
Db  6978 CAAGAUUGUGUGCCUCACCCCAAAUACACAUCCACCCGGAGAAUUAACAUCAUGCUUCAAAU 7037

Qy  8366 GTTACCCTCGGAATTACCAAAAGTGGCAGCCTTACTTTCTTACAGAGATCTTCGT 8425
         |:|||||||:|:|||||::::|||||||:||:|||:||:||||::|||||:|::|
Db  7038 GUUACCCUGGAAUUACAGCCAAAAGUGGGCAAGCCUUACUCUACUUUCUUACAAGAGAUCUUCCGU 7097

Qy  8426 ATCCCCCTTGGCAGGTGCTCTGCCGAGGGTCTGGATACAACCCCAGTGCTGCTGGATT 8485
         |:|||:|||||:|:|:|::|||||||:||||:||:|||||||||:||:|||||||||:
Db  7098 AUCCCCCUUGGCCAGCGUCCUCGGCGUCCGAGGGUCUGGAUACAACCCCAGUGCUGCUGGAUU 7157

Qy  8486 GGGTATCTAATACATCACTACCCATGTTTGTGGTTAGCGTGTTGGCTGTCCATTTC 8545
         |||||:|||:||||||:||||:||||::|:|||||||||||||||||||:||||||
Db  7158 GGGUAUCUAAUACAUCAUCAUCCCAUGUUUGUGGUUAGCGUGUUGGCUGUCCAUUUC 7217

Qy  8546 ATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGAGACTGTGACCTTTGACTGGTATGGG 8605
         |:|||||||||||||||||||||||||||:|:::||:|||:||:||||||||||||||||
Db  7218 AUGGAGCAGAUGCUCUUUGAGGACAAACUUCCCGAGACUGUGACCUUUGACUGGUAUGGG 7277

Qy  8606 AAAAATTATACGGTGCCTGTAGAAGATCTGCCCAGCATCATTGCTGCTGTGCACGGTATT 8665
         ||||||||| ||:||:|:|:|:|||:||||||:||:||:|||:||:||||||||||||||
Db  7278 AAAAAUAUACGGUGCCGUGUAGAAGAUCUGCCCAGCAUCAUUGCUGCUGUGCACGGUAUU 7337

Qy  8666 GAGGCTTTCTCGGTGGTGCGCTACACCAAACGTGAGATCCTCAGAGTTTCCAATCACTA 8725
         |||||||||:|:|||:|:|||||||||::|||:|||||:||||||||||:|:|||||||
Db  7338 GAGGCUUUCUCGGUGGUGCGCUACACCAAACGUGAGAUCCUCAGAGUUUCCAAUCACUA 7397

Qy  8726 ACAGACATGACCATGCCCCCCTCGCGAGCCCCTGGCGAGTCCTCGCC 8785
         ||||||||||||||||||||||||:|||||:|:|||||||||:|:||
Db  7398 ACAGACAUGACCAUGCCCCUGCGAGCCCUGCGAGUCCUGGCGAGUCCUCGCC 7457

Qy  8786 AGCGCCAAGAGGCGCTGGCGAGCCACCCAAATTGGCTCGCTTCCTTCTTCGGCATGCT 8845
         ||||||||||||||||:||||:|||||||||:|||||:||::|||:||::|:||:|||
Db  7458 AGCGCCAAGAGGCGCGUGGCGAGCCACGCAAAAUUGGCUCGCUUCCUUCUUCUGGCAUGCU 7517

Qy  8846 ACATCTAGAACCTTCTACCAGATTGGATAAGACGAGCCTGGCTCGGTACCACACTTTCAAT 8905
         ||||||:|||||:||||:|:||||||:|||||:||||:||||:||||||||||:|||:|||
Db  7518 ACAUCUAGAACCUUCUACCAGAUUGGAUAAGACGAGCCUGGCUCGGUACCACACUUUCAAU 7577

Qy  8906 TATTGTGATGTTTACTCCCCGGAGCGGGATGTCTTTATTACACCACAGAGAGATTGCAG 8965
         :|:::|:||:|||||::|:::||||||||:|:|||||:::::|||||||||||:||||||
Db  7578 UAUUGUGAUGUUUACUCCCCGGAGCGGGAUGUCUUUAUUACACCACAGAGAGAUUGCAG 7637

Qy  8966 AAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTTTGCCCTAGGGCTCATTGCTGTTGA 9025
         |||||:::|||:|||:|:||||||||||||||:|||:||:|||||:|||||||:|||||
Db  7638 AAGUUCCUUGUGAAGUAUUUGGCUGUCAUUGUUUUGCCCUAGGGCUCAUUGCUGUUGGA 7697

Qy  9026 TTAGCCATGAGTCAGCTGAACCTGAAACCCCCAAATTCAAAATTAACACAGTTTTTTTTTTTTTT 9085
         ::|||:||||:||:||||||||::||||||||:||||||||||:||:|:||||||||||||||||
Db  7698 UUAGCCAUCAGCUGAACCUCGGUGAACCCCCAAAUUCAAAAUUAACUAACAG--UUUUUUUUUUUUU 7755

Qy  9086 TTTTTAGGCGACGCGCAACAGGGGAGCACCAGGGGAGACCCCGGGCTTAACGACCCCGGA 9145
         :::::|||:|:|:|||||||||||||||:|||||||||:|||:||||:|||||||||:||
Db  7756 UUUUUAGGGCAGCGCGCAACAGGGGAGCACCAGGGGAGACCCCGGGCUUAACGACCCCGGA 7815

Qy  9146 GGGCGACCATGGTCGGATCAGAAGCCGTTTCGGGTGTCGAAGGCCATGGGATGACCTC 9205
         ||||:|||:|||||||||:||||:|:|||:||||||:|||||:|:|||||||||||||
Db  7816 GGGCGACCAUGGUCGGAUCAGAAGCCGUUUCGGGUGUCGAAGGCCAUGGGAUGAGACGUC 7875

Qy  9206 CCTTCTGGCTCATCCACAAAAACCGTCTTCGGGGTGTGGCTCTGGCTTGTGTGGAA 9265
         ::|||||||:|||||::|:||||||:||:||||||:|||::|||||:||||:||||
Db  7876 CCUUCUGGCCUCAUCCACCACAAAAACCGUCUUCGGGGUGUGGCUCUGGCUUGUGUGGAA 7935

Qy  9266 GCAGTCAGTATATTCCCGTGTGTGGTGCGACGGCCTCACGACGTATTGTCCGGCTGC 9325
         ||||:|:|||:|||||||||||||:||:||||:|||||||||:||:||:||||::||
Db  7936 GCAGUCAGUAUAAUUCCCGUGUGUGGUGCGACGGCCGCCUCACGACGUACUGUCCGGCUGC 7995

Qy  9326 AGAGCGGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCCAAGCGGAGGGCAACCCCCGCT 9385
```

CC   phage HGBV cDNA library. The cDNA clone AAT00052, which encodes the
CC   proteins AAR82066-71 (the 6 possible reading frames), was rescued from
CC   the lambda phage, searched against a sequence database and found to be an
CC   unique HGBV sequence. Reagents which comprise the HGBV DNA, or its
CC   protein prods. can be used for the diagnosis, therapy or in a vaccine to
CC   prevent HGBV infection.. (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 4268 BP; 1036 A; 1063 C; 1063 G; 1086 T; 0 U; 20 Other;

Query Match   44.2%; Score 4159; DB 2; Length 4268;
Best Local Similarity   98.1%; Pred. No. 0;
Matches 4215; Conservative   20; Mismatches   30; Indels   30; Gaps   2;

Qy   3456 TGGCTCATCCCACAGGCTCATACACCGTTGACCGGCTAATGACCAGGACA 3515
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TGGCTCATCCCACAGGCTCCATACACCGTTGACGCGGCTAATGACCAGGACA 60

Qy   3516 TCTATCAACCACCATGTGGAGCTGGGTCCCTTACTCGTGTCTTGCGGTGGTG 3575
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TCTATCAACCACCATGTGGAGCTGGGTCCCTTACTCGTGTCTTGCGGTGGTG 120

Qy   3576 GGTATCTGGTAACACGACTGGGGTCATTGGTTGAGGTCAACAAATCCGATGA 3635
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGTATCTGGTAACACGACTGGGGTCATTGGTTGAGGTCAACAAATCCGATGA 180

Qy   3636 GGTGTGTGTGCGGGGCCCTTCCCATGGCTGTTGCCAAGGGTTCTTCAGGTCG 3695
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GGTGTGTGTGCGGGGCCCTTCCCATGGCTGTTGCCAAGGGTTCTTCAGGTCG 240

Qy   3696 TGTGGTTCCTCCCGGCATGTTATTGGGATGTTCACCGTGCTGTAGAAATTCTCG 3755
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGTGGTTCCTCCCGGCATGTTATTGGGATGTTCACCGTGCTGTAGAAATTCTCG 300

Qy   3756 TCAGTCAGATTAGGGTTAGGCCGTTGGTGTGCTGGATGGTGTGTGTGCTGGATA 3815
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TCGGCCAGATTAGGGTTAGGCCGTTGGTGTGCTGGATGGTGTGTGTGCTGGATA 360

Qy   3816 ATGCCACTCTGATACAAAACTTGCCTAAGGTATTCAGTGCAAATTTTAATTG 3875
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ATGCCACTCTGATACAAAACTTGCCTAAGGTATTCAGTGCAAATTTTAATTG 420

Qy   3876 CCCCCCACTGGAGCGGCAAGTCAACCAAATTACCACTTTCTTACATGGAGAAGTATG 3935
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CCCCCACTGGAGCGGCAAGTCAACCAAATTACCACTTTCTTACATGGGAGAAGTATG 480

Qy   3936 AGGTCTTGGTCCTAAATCCCAGTGTGGCTACAACAGCATCAATGCCAAAGTACATGCACG 3995
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AGGTCTTGGTCCTAAATCCCAGTGTGGCTACAACAGCATCAATGCCAAAGTACATGCACG 540

Qy   3996 CGACGTACGGCGTGAATCCGTGATATTTAATGGCAAATGTACAACACACAGGGGCTT 4055
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CGACGTACGGCGTGAATCCGTGATATTTAATGGCAAATGTACAACACACAGGGGCTT 600

Qy   4056 CACTTACGTACGACACATATGGCATGTCACCGGACATGTTCCCGGACTATGATGT 4115
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CACTTACGTACGACACATATGGCATGTCACCGGACGATGTTCCCGGACTATGATG 660

Qy   4116 TAATCATTTGTGACGAATGCCATGTCGACCACCGTGTTGGGCATTGGGAAAGG 4175
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TAATCATTTGTGACGAATGCCATGTCGACCACCGTGTTGGGCATTGGGAAAGG 720

Qy   4176 TCCTAACCGAAGCTCCATCCAAAAATGTAGGTGGTTGGTTCTTGCCACGGCTACCCCC 4235
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TCCTAACCGAAGCTCCATCCAAAAATGTAGGGTCAGTGGTTCTTGCCACGGCTACCCCCC 780

Qy   4236 CTGGGAGTAATCCCTACACCACATACCGCCAACATTAACCGATGAAGGCA 4295
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CTGGGAGTAATCCCTACACCACATACCGCCAACATAACCTGATGAAGGCA 840

Qy   4296 CTATCCCCTTTCATGGAAAAAGATTAAGGAGGAAAAATCTGAAGAAGGGGAGACACCTTA 4355
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CTATCCCCTTTCATGGAAAAAGATTAAGGAGGAAAAATCTGAAGAAGGGGAGACACCTTA 900

Qy   4356 TCTTTGAGGCTACCAAAAAACACTGTGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAA 4415
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCTTTGAGGCTACCAAAAAACACTGTGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAA 960

Qy   4416 TAACAGCTGTCTTCTTACTATAGGGGATGGATGTGACATCTCAAAAATCCCTGAGGGGGACTGTG 4475
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TAACAGCTGTCTTCTTACTATAGGGGATGGATGTGACATCTCAAAAATGCTGAGGGGGACTGTG 1020

Qy   4476 TAGTAGTTGCCACTGATGCCTTGTGTACAGGGTACACTGGTGACTTTGATTCCGTGTATG 4535
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TAGTAGTTGCCACTGATGCCTTGTGTACAGGGTACACTGGTGACTTTGATTCCGTGTATG 1080

Qy   4536 ACTGCAGCCTCATGGTAGAAGGCACATGCCATGTGACCTTGACCCTGACCTACTTCACCATGG 4595
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ACTGCAGCCTCATGGTAGAAGGCACATGCCATGTGACCTTGACCTTGACCTACTTCACCATGG 1140

Qy   4596 GTGTTCGTGTGTGCGGGGTTCAGCAATAGTAGTGACCAGGCCAGGCTAGGGGCCACAGGCC 4655
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GTGTTCGTGTGTGCGGGGTTCAGCAATAGTAAAGGCCACCAGGCTAGGGGCCGCACAGGCC 1200

Qy   4656 GTGGGAGAGCTGGTAACACGAGTTGTACCCCTTCGGGTATGGTTC 4715
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GTGGGAGAGCTGGTAACACGAGTTGTACCCCTTCGGGTATGGTTC 1260

Qy   4716 CTGAATGCAACATTGTTGAAGCCTTGAAGGTCATGGTATGGTTGTCATCAA 4775
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CTGAATGCAACATTGTTGAAGCCTTGAAGGTCATGGTATGGTTGTCATCAA 1320

Qy   4776 CAGAAGCTCAAACTATTCTGGACACCTATCGGACTGGGTTACCTGCGATAGGAG 4835
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CAGAAGCTCAAACTATTCTGGACACCTATCGGACTGGGTTACCTGCGATAGGAG 1380

Qy   4836 CAAATTTGGACGAGTGGGCTGATCTCTTTCTATGGTCAACCCGAACCTCATTTGTCA 4895
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CAAATTTGGACGAGTGGGCTGATCTCTTTCTATGGTCAACCCGAACCTCATTTGTCA 1440

Qy   4896 ATACTGCAAAAAGAAACTGCTGACAATTATGTTTGTTGACTGGCAGCCCAACTACAACTGT 4955
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 ATACTGCAAAAAGAAACTGCTGACAATTATGTTTGTTGACTGGCAGCCCAACTACAACTGT 1500

Qy   4956 GTCATCAGTATGGCTATGCTGCTCCAATGACGGCACCACGCGGTGGCAGGGAGCCCGGCTTG 5015
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GTCATCAGTATGGCTATGCTGCTCCCATGACGGCACCACGCGGTGGCAGGGAGCCCGGCTTG 1560

Qy   5016 GGAAAAAACCTTGTGGGGTTCTGTGGCCTTGACGCCGGCTGCGACGGTGTCCTCGCCCAG 5075
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GGAAAAAACCTTGTGGGGTTCTGTGGCCTTGACGCCGGCTGTCGACGGTGTCCTCGCCCAG 1620

Qy   5076 AGCCCAGCGAGTGACCAGATACCCAAATGTGCTTCACTGAAGTCAATACTTCTGGGACAG 5135
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 AGCCCAGCGAGTGACCAGATACCCAAATGTGCTTCACTGAAGTCAATACTTCTGGGACAG 1680

Qy   5136 CCGCACTCGCTCGTGTGGCGTTGGGCGTTGCTGTTGGCCTATCGTAGCCATTGACACTTTGGCG 5195
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 CCGCACTCGCTCGTGTGGCGTTGGGCGTTGCTGTTGGCCTATCGTAGCCATTGACACTTTTGGGCG 1740

Qy   5196 CCACTTGGTGACGAAGAAGAAATCGTGGAGGAGTGCATCATTCCGTGTGCTACTGTCGCCC 5255
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 CCACTTGGTGACGAAGAAGAAATCGTGGAGGAGTGCATCATTCCGTGTGCTACTGTCGCCC 1800

Qy   5256 CAGTGGTTGCAATTGACAAGCTGAAGAGAGAAATCGTGAAGAGTACAATCGTGCATCCAGGGCCA 5315
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CAGTGGTTGCAATTGACAAGCTGAAGAGAGAAATCGTGGAGAGTGCAATCGTACTAGTCCTTCAGATTGG 1860

Qy   5316 TGGTTGCTGCAATTGACAAGCTGCCAACATAACTGATGATTCAATTAACCGATGAGGCA 5375
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGGTTGCTGCAATTGACAAGCTGAAGAGTACAATCAATCAGTACATTACCACATTGG 1920

Qy   5376 AAACCGCCCTTGAAAAACTTAACACCTTTCTTGGGCCTTCATGCAGCCTACAAATCCTTGCTA 5435
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AAACCGCCCTTGAAAAACTTAACACCTTTCTTGGGCCTTCATGCAGCCTACAAATCCTTGCTA 1980

Qy   5436 TCATAGAGTATTGCTGTGTGGTTTAGTCACTTTACCTGACAATCCCTTTGCATCATGCGTGT 5495

```
Db                  ||||||||||||||||                |||| |||||||||||||||||||||||||||||||||
     1981           TCATAGAGTATTGCTGTGGCTTAGTCACTTTACCTGACAATCCCTTTGCATCATGCGTGT  2040

Qy                                                  |||||||||||||||||||||||||||||||||||
     5496           TTGCTTTCATTGCGCGGTATTACTACCCCACTACCTCACCCCGTTGACGGCGCACTGGGGT  5555
Db                  ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
     2041           TTGCTTTCATTGCGCGGTATTACTACCCCACTACCTCACCCCGTTGACGGCGCACTGGGGT  2100

Qy                                                  |||||||||||||||||||||||||||||
     5556           TATTTGGAGGCGCAATTGCGTCCAAGCTTACAGACGTTAGAGGCGCACTGGCGTTCATGA  5615
Db                  |||||||||||||||||||||||||||||||||||||||||||||
     2101           TATTTGGAGGCGCAATTGCGTCCAAGCTTACAGACGTTAGAGRCGCACTGGCGTTCATGA  2160

Qy                                                  ||||||||||||||||||||||||||||
     5616           TGGCCGGGGCGTGCGGGAACAGCTCCTTGGTACATGGACTCGGTGGGTTTGTCTTTGACA  5675
Db                  ||| ||:|||||||||| |  |||||||||||||||||||||||||||||||||||||
     2161           TGGCCGGGGCGTGYGGGAACAGCATGCATGGACTCGGTGGGTTTGTCTTTGACA  2220

Qy                                                  ||||||||||||||||||||||||||||||
     5676           TGCTAGGCGGCTATGCTGCCGCCTCATCCACTGCTTGCTTGTTGACATTTAAATGCTTGATGG  5735
Db                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
     2221           TGCTAGGCGGCTATGCTGCCGCCTCATCCACTGCTTGCTTGTTGACATTTAAATGCTTGATGG  2280

Qy                                                  |||||||||||||||||||||||||||||
     5736           GTGAGTGTGCCCACTATGGATCAGCTTGCTGGTTTAGTCTACTTCCGCGTTCAATCCGCCG  5795
Db                  |||||||||||||||||||||||||||||||||||||||||||||||||
     2281           GTGAGTGGCYCACTATGGATCAGCTTGCTGGTTTAGTCTACTTCCGCGTTCAATCCGCCG  2340

Qy                                                  |||||||||||||||||||||||||||||
     5796           CAGGAGTTGTGGGCGTCTTGTCAGCTTGTCTGTGCAATGTTTGCTTTGACAACAGCAGGGGCCAG  5855
Db                  |||||||||||||||||||||||||||||||||||||||||||||||||
     2341           CAGGAGTTGTGGGCGTCTTGTCAGCTTGTCTGTGCAATGTTTGCTTTGACAACAGCAGGGGCCAG  2400

Qy                                                  |||||||||||||||||||||||||||||
     5856           ATCACTGCCCACAGCAGACTTCTTACTATGCTTGCTAGGACAACACTGTATGTAATGAGT  5915
Db                  ||||||||||||||||||||||||||| : ||||||||||||||||||||||||||
     2401           ATCACTGCCCACAGCAGACTTCTTCTTCAGCTTGCTAGGACAACACTGTATGTARTGAGT  2460

Qy                                                  |||||||||||||||||||||||||||||
     5916           ACTTTATTGCCACTCGTTGACATCCGCAGGAAGATACTTGGGCATTCTGGGGCATCTACCC  5975
Db                  ||||||||||||||||||||||||||||||||||||||||||||||||||
     2461           ACTTTATTGCCACTCGTGACATCCGCAGGAAGATACTTGGGCATTCTGGGGCATCTACCC  2520

Qy                                                  |||||||||||||||||||||||||||||
     5976           CCTGGAGTTGTCATATCATCAGCTTGCATCCGTTGGCTCCACACCCCGACGGAGATGATTGCG  6035
Db                  ||||||||:|||||||||||||||||||||||: |||||||||||||||||||||||
     2521           CCTGGAGTRTCATATCATCAGCTTGGCTYCACACCCCGACGGAGATGATTGCG  2580

Qy                                                  |||||||||||||||||||||||||||||
     6036           GCCTCATTGCTTGGGGTCTAGAGATTTGGCAGTATGTGCAGTTTTGTTGAGCAATTTCTTGTTGCT  6095
Db                  ||||||||||||||||:|| |||||||||||||||||||||||||||||||||||
     2581           GCCTCATTGCTTGGGGTCTARAGATTTGGCAGTATGTGCAGTTTTGTTGAGCAATTTCTTGTTGCT  2640

Qy                                                  |||||||||||||||||||||||||||||
     6096           TTAAATGTCCTTAAAGCTGGAGTTCAGAGCATGGTTAACATTCCTGGTTGTCCTTTTCTACA  6155
Db                  |||||||||||||||||||||||||||||||||||||||||||||||||
     2641           TTAATGTCCTTAAAGCTGGAGTTCAGAGCATGGTTAACATTCCTGGTTGTCCTTTTCTACA  2700

Qy                                                  |||||||||||||||||||||||||||||
     6156           GCGCTAGAAGGGGTACAAGGGGGTACACAGTGATTGGACTCAGGTATGCTCCAAGCACGCTGTC  6215
Db                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
     2701           GCGCTAGACCGGACCCGGACCAACTGATTGGACTCAGGTATGCTCCAAGCACGCTGTC  2760

Qy                                                  |||||||||||||||||||||||||||||
     6216           CATGGCGGTGCTGAACTCATCTTTTCGTTGAGAATGGTTTGCAAAACTTTACAAAGGAC  6275
Db                  |||||||||||||||||||||||||||||||||||||||||||||||||
     2761           CATGGCGGTGCTGAACTCATCTTTTCGTTGAGAATGGTTTGCAAAACTTTACAAAGGAC  2820

Qy                                                  |||||||||||||||||||||||||||||
     6276           CCAGAACTGTTCAGTCAAATTACTGGAGTTCAGAGCATGCATGGTGTGTGGGT  6335
Db                  |||||||||||||||||||||||||||||||||||||||||||||||||
     2821           CCAGAACTGTTCAGTCAAATTACTGGAGAGGGCTGTTCCAGTCAGTGCTGTGTGGGT  2880

Qy                                                  |||||||||||||||||||||||||||||
     6336           CGGCTAGACCGGACCCAACTGATTGGACTCAGGTATGCTTGTCGTCAATTATGGCGTTAGGGACT  6395
Db                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
     2881           CGGCTAGACCGGACCCAACTGATTGGACTCAGGTATGCTTGTCGTCAATTATGGCGTTAGGGACT  2940

Qy                                                  |||||||||||||||||||||||||||||
     6396           ACTGTAAAATATGAGAAAATGGGAGATCACATTTTGTTACAGCGAGTATCCTCTCCAAATG  6455
Db                  |||||||||||||||||||||||||||||||||||||||||||||||||
     2941           ACTGTAAATATGAGAAAATTGGGAGATCACATTTTGTTACAGCGAGTATCCTCTCCAAATG  3000

Qy                                                  |||||||||||||||||||||||||||||
     6456           TCTGTTTCACCCAGGTGCCCCCAACCTTGAGAGCTGCAGTGGCCGTGAGCACGGCGTACAGG  6515
Db                  ||||||||||||||||||||||||||||||||||||||||||||||||||
     3001           TCTGTTTCACCCAGGTGCCCCCAACCTTGAGAGCTGCAGTGGCCGTGAGCACGGCGTACAGG  3060

Qy                                                  ||||||:||||||||||||||||||||||
     6516           TTCAGTGTTATCGGTGAGCCCAAAACTCCTTGGACGACAATCCCTTGCTTGCTGTTACGGTC  6575
```

```
Db                  |||| ||||||||||||||||||||||||||||||||||||||||||||
     3061           TTCAGYGTTATCTAGGTGAGCCCAAAACTCCTTGCTTGCTGTTACGGTC  3120

Qy                                                  ||||||||||||||||||||||||||||||
     6576           CTGACGGGTAAGGGTAAAACTGTTAAGCTTCCCTTCCGCGGTTGACGGCGTCACACACTGGTG  6635
Db                  ||||||||||||||||||||||||||||||||||||||:|||||||
     3121           CTGACGGGTAAGGGTAAAACTGTTAAGCTTCCCTTCCGCGGTTGACGGCGCACACACTGGTG  3180

Qy                                                  |||||||||||||||||||||||||||||
     6636           TGGCCATGCAACTTAATTTGCGTGATGCAGTCACTTGAGACAAATGACTGTAATTCCACAAACA  6695
Db                  |||||||||||||||||||||||||||||||||||||:|||||||:|||||||||
     3181           GTGCGGATGCAACTTAATTTGCGTGATGCAGTCACTTGAGTCAAATGACTGTAATTCCATAAACA  3240

Qy                                                  |||||||||||||||||||||||||||||
     6696           ACACTCCTAGTGATGAAGCCGCAGTGTCCGCTCTTGTTTCAAACAGGAGTTGCCGCGTA  6755
Db                  ||||||||||||||||||||||||||||||||||||||||||||||||
     3241           ACACTCCTAGTGATGAAGCCGCAGTGTCCGCTCTTGTTTTCAAACAGGAGTTGCGCGGCGTA  3300

Qy                                                  |||||||||||||||||||||||||||||
     6756           CAAACCAATTGCTTGAGCAATTTCAGCTGGCGGTGCACCACCAAAACTGCCAGCCCCCCT  6815
Db                  ||||||||||||||||||||||||||||||||||||||||||||||||
     3301           CAAACCAATTGCTTGAGCAATTTCAGCTGGCGGTTGACACCACCAAAACTGCCAGCCCCCCT  3360

Qy                                                  ||  |||||||||||||||||||||||||
     6816           CC--ATCCAAGAGGTAGTGGTAAGAAAGCGCCAGTTCCGGCAAGAACTGGTTCGCTTA  6872
Db                  ||||||||||||||||||||||||||||||||||||||||||||
     3361           CCCAGATCGAAGAGGTAGTGGTAAGAAGCGCCAGTTCCGGCAAGAACTGGTTCGCTTA  3420

Qy                                                  |||||||||||||||||||||||||||||
     6873           CCTTGCCTCCCCCTCCAGGAGTCTTCAAACCTCCCTCCTGAAAGGTCATGTCTGTTGAAAGCCTGAACGAA  6932
Db                  ||||||||||||||||||:||||||:||||||||||||||||||||||
     3421           CCTTGCCTCCCCCTCCGGAGTCTTCAAACCTCCCTCCTGAAAGGTCATGTCTGTTGAAAGCCTGAACGAA  3480

Qy                                                  |||||||||||||||||||||||||||||
     6933           GTGACCCGTAGAAGGTCTTCAAACCTCCCTCCTTCACCACCTGTTCTCACAGTGTGGCCA  6992
Db                  |||||||||||||||||:||:|||||||:|||||:|||||||:|||||R|||||||||
     3481           GTGACCCGTTAGAAGGTCTTCAAMCCTCCCTCCTTCACCACCTGTTCTCCAGTRGCCA  3540

Qy                                                  |||||||||||||||||||||||||||||
     6993           TGCCGATGCCCCTGTTGGGAGCGGGTGAGTGTAACCCTTCACTGCAATTGGATGTGCAA  7052
Db                  ||||||||||||||||||||||||||||||||||||||||||||||||
     3541           TGCCGATGCCCCTGTTGGGAGCAGGTGAGTGTAACCCTTCACTGCAATTGGATGTGCAA  3600

Qy                                                  |||||||||||||||||||||||||||||
     7053           TGACCGAAACAGGCGGAGCGCCCTGATGATTACCCGAGTGATTACCCCTCCCAAAAAGGAGGTCT  7112
Db                  ||||||||||||:||||||||||||||||:::|||AK|||||||||||||||||||||
     3601           TGACCGAAACARGYGGAGKCCCCWSAKRATTACCCGAGTGATTACCCCTCCCAAAAAGGAGGTCT  3660

Qy                                                  |||||||||||||||||||||||||||||
     7113           CTGAATGGTCAGACGAAGTTGGTCGACGGCTACAACGGCTTCAGCTACGTTACTGGCC  7172
Db                  ||||||||||||||||||||||||||||||||||||||||||||||||
     3661           CTGAATGGTCAGACGAAGTTGGTCAACGACTACAACGGCTTCAGCTACGTTACTGGCC  3720

Qy                                                  |||||||||||||||||||||||||||||
     7173           CCCCGTACCCTAAGATACGGGGAAAGGATTCCACTCAGTCAGCCCCGCCAAACGGCCTA  7232
Db                  ||||||||||||||||||||||||||||||||||||||||||||||||
     3721           CCCCGTACCCTAAGATACGGGGAAAGGATTCCAATCAGCCACCGCCACCCGCCAAACGGCCTA  3780

Qy                                                  |||||||||||||||||||||||||||||
     7233           CAAAAAAGAAGTTGGGAAAGAGTGAGTTTTCGTGCAGCATGAGTCACACCTGGACCGACG  7292
Db                  ||||||||||||||||||||||||||||||||||||||||||||||||
     3781           CAAAAAAGAAGTTGGGAAAGAGTGAGTTTTCGTGCAGCATGAGTCACACTTGGACCGACG  3840

Qy                                                  |||||||||||||||||||||||||||||
     7293           TGATTAGCTTCAAAACTGCTTCTAAAGTTGTCGTGTCTCGGGCCATCACTAGTGGTT  7352
Db                  |||||||||||||||||||||||||||||||||||||||||||||
     3841           TGATTAGCTTCAAAACTGCTTCTAAAGTTGTCGTGTCTCGGGCCATCACTAGTGGTT  3900

Qy                                                  |||||||||||||||||||||||||||||
     7353           TCCTCAAACAAAGATCATTGGTGTGTATGTGACTGGAGCGGGGATGCGGAGCTTAGAAAAAC  7412
Db                  |||||||||||||||||||||||||||||||||||||||||||||||
     3901           TCCTCAAACAAAGATCATTGGTGTGTATGTGACTGGAGCGGGGATGCGGAGCTTAGAAAAAC  3960

Qy                                                  |||||||||||||||||||||||||||||
     7413           AAAAAGTCACTATTAATAGACAACCTCTGTTCCCCCATCATCACCACACAAGCAAGTGAGAT  7472
Db                  |||||||||||||||||||||||||||||||||||||||||||||||||
     3961           AAAAAGTCACTATTAATAGACAACCTCTGTTCCCCCATCATCACCACACAAGCAAGTGAGAT  4020

Qy                                                  |||||||||||||||||||||||||||||
     7473           TGGGCTAAGGAAAAAGCTTCAAAAGCTTCAAAAGTTGTCGTGTCTGGGACTATGTGGGACTATGAGTAGCAG  7532
Db                  |||||||||||||||||||||||||||||||||||||||||||||||
     4021           TGGCTAAGGAAAAAGCTTCAAAAGCTTCAAAAGTTGTCGTGTCTGGGACTATGTGGGACTATGAGTAGCAG  4080

Qy                                                  |||||||||||||||||||||||||||||
     7533           CTCACACGCCCTTGCTAAGTCTGCTAAGTCCCACATCACTCGGCCTTGGGCGGCCACTCGATGTTC  7592
Db                  |||||||||||||||:||||||||||||||||||||||||-----||||||
     4081           CTCACACCGCCCTCGTCTAAGTCTGCTAAGTCCCACTCGGCCTTCGGGCGGCCACTGA  4135

Qy                                                  |||||||||||||||||||||||||||||
     7593           GTTCTGGAGCAGCCCGGCAAGGCTGTTCTGAGACTTGCAGAAGTGTGTCGAGGCAGGTGAGA  7652
Db                                                    -------TGTTCTCGGACTTGCAGAAGTGTGTCGAGGCAGGTGAGA  4173
```

```
Qy    7653  TACCGAGTCATTATCGGCAAACTGTGATAGTTCCAAAGGAGGAGGTCTTCGTGAAGACCC 7712
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4174  TACCGAGTCATTATCGGCAAACTGTGATAGTTCCAAAGGAGGAGGTCTTCGTGAAGACCC 4233

Qy    7713  CCCAGAAACCAACAAAGAAACCCCAGGCTTATC 7747
            |||||||||||||||||||||||||||||||||
Db    4234  CCCAGAAACCAACAACAAAGAAACCCCAGGCTTATC 4268

RESULT 14
AAA55298
ID    AAA55298 standard; DNA; 4268 BP.
XX
AC    AAA55298;
XX
DT    06-AUG-2003 (revised)
DT    30-AUG-2000 (first entry)
XX
DE    Hepatitis GB virus nculeotide sequence SEQ ID NO:80.
XX
KW    Hepatitis GB virus; HGBV; diagnosis; therapeutic; immunogenic; infection;
KW    detection; characterisation; hepatitis; ss.
XX
OS    Hepatitis GB virus.
XX
PN    US6051374-A.
XX
PD    18-APR-2000.
XX
PF    07-JUN-1995;       95US-00488445.
XX
PR    14-FEB-1994;       94US-00196030.
PR    13-MAY-1994;       94US-00242654.
PR    29-JUL-1994;       94US-00283314.
PR    23-NOV-1994;       94US-00344185.
PR    23-NOV-1994;       94US-00344190.
PR    30-JAN-1995;       95US-00377557.
XX
PA    (ABBO ) ABBOTT LAB.
XX
PI    Dawson GJ, Leary TP, Muerhoff AS, Pilot-Matias TJ, Buijk SL;
PI    Mushahwar IK, Simons JN, Desai SM, Erker JC, Schlauder GG;
XX
DR    WPI; 2000-338307/29.
XX
PT    Detecting target hepatitis GB virus nucleic acid in a test sample
PT    suspected of containing HGBV comprises reacting the test sample the HGBV
PT    polynucleotide probe and detecting the complex that contains target HGBV.
XX
PS    Example 5; Col 205-210; 369pp; English.
XX
CC    The present invention describe a method for detecting target hepatitis GB
CC    virus (HGBV) nucleic acid (THN) in a test sample (T) suspected of
CC    containing HGBV. The method involves reacting (T) with a HGBV
CC    polynucleotide probe (I) containing 15 contiguous nucleotides, and which
CC    selectively hybridises to the HGBV genome or its full complement, and
CC    detecting the complex that contains THN, indicating the presence of
CC    target HGBV. The method is used for detecting target HGBV nucleic acid in
CC    the test sample suspected of containing HGBV and for characterisation of
CC    newly ascertained etiological agent of non-A, non-B, non-C, non-D and non
CC    -E hepatitis causing agents collectively termed as hepatitis GB virus.
CC    AAA55270 to AAA55489 and AAB09985 to AAB09480 represent nucleotide and
CC    protein sequences used in the exemplification of the present invention.
CC    (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ    Sequence 4268 BP; 1036 A; 1063 C; 1063 G; 1086 T; 0 U; 20 Other;

Query Match       44.2%;   Score 4159;   DB 3;   Length 4268;
Best Local Similarity   98.1%;   Pred. No. 0;
Matches 4215;   Mismatches 20;   Conservative 0;   Indels 30;   Gaps 2;

Qy    3456  TGGCTCATCCCACAGGCTCTATACACCCAATAACCGTTGACGCGGCTAATGACCAGGACA 3515
```

```
Db       1  TGGCTCATCCCACAGGCTCTATACACCCCAATAACCGTTGACGCGGCTAATGACCAGGACA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    3516  TCTATCAACCACCATGTGGAGCTGGGTCCCTTACTCGGTGCTCTTGCGGGGGAGACCAAGG 3575
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61  TCTATCAACCACCATGTGGAGCTGGGTCCCTTACTCGGTGCTCTTGCGGGGGAGACCAAGG 120

Qy    3576  GGTATCTGGTAACAGACTGGGTCAGGGTCATTGGTTGAGGTCAACAAATCCGATGACCCTTATT 3635
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121  GGTATCTGGTAACAGACTGGGTCATTGGTTGAGGTCAACAAATCCGATGACCCTTATT 180

Qy    3636  GGTGTGTGTGCGGGGCCCTTCCCATGGCTGTGTTGCCAAGGGTTCTTCAGTGTGCCCGATTC 3695
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181  GGTGTGTGTGCGGGGCCCTTCCCATGGCTGTGTTGCCAAGGGTTCTTCAGTGTGCCCGATTC 240

Qy    3696  TGTGCTCCTCCGGGCATGTTATTGGGAATGTTCACCGGTGCTAGAAATCTCGGCGGTTCAG 3755
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241  TGTGCTCCTCCGGGCATGTTATTGGGAATGTTCACCGGTGCTAGAAATCTCGGCGGTTCAG 300

Qy    3756  TCAGTCAGATTAGGGTTAGGCCGTTGGTGTGTGCTGGATACCATCCCAGTACACAGCAC 3815
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301  TCGGCCAGATTAGGGTTAGGCCGTTGGTGTGTGCTGGATACCATCCCAGTACCACAGCAC 360

Qy    3816  ATGCCACTCTTGATACAAAACCTACTGTGCCTAACGAGTATTCAGTGCAAATTTTAATTG 3875
            |||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db     361  ATGCCACTCTTGATACAAAACCTACTGTGCCTAACGAGTATTCAGTGCAAATTTTAATTG 420

Qy    3876  CCCCCACTGGCAGCGCGCAAGTACCAACCAAATTACCACTTTCTTACATGCAGGAGAAGTATG 3935
            ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db     421  CCCCCACTGGCAGCGCGCAAGTACCAACCAAATTACCACTTTCTTACATGCAGGRGAAGYATG 480

Qy    3936  AGGTCTTGGTCCTAAATCCCAGTGTGGCTACAACAGCATCAATGCCAAAGTACATGCACG 3995
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481  AGGTCTTGGTCCTAAATCCCAGTGTGGCTACAACAGCATCAATGCCAAAGTACATGCACG 540

Qy    3996  CGACGTACGGCGTGAATCCAAATGCTATTTTAATGGCAAATGTACCAACACAGGGGCTT 4055
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541  CGACGTACGGCGTGAATCCAAATGCTATTTTAATGGCAAATGTACCAACACAGGGGCTT 600

Qy    4056  CACTTACGTACAGCACACATATGGCATGTACCTGACCGGAGCACTGTTCCCGGAACTATGATG 4115
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601  CACTTACGTACAGCACACATATGGCATGTACCTGACCGGAGCACTGTTCCCGGAACTATGATG 660

Qy    4116  TAATCATTGTGACGAATGCCATGCTACCAGCACCGTGTTGGGCATTGGAAAGG 4175
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661  TAATCATTGTGACGAATGCCATGCTACCAGCACCGTGTTGGGCATTGGAAAGG 720

Qy    4176  TCCTAACCGAAGCTCCATCCAAAAATGTTAGGCTAGCTAGTTGGTTCTTGCCACGGCTACCCCC 4235
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721  TCCTAACCGAAGCTCCATCCAAAAAATGTTAGGCTAGCTAGTTGGTTCTTGCCACGGCTACCCCC 780

Qy    4236  CTGGAGTAATCCTACCACCACATGCCAACATAACTGAGATTCAATTAACGATGAAGGCA 4295
            ||||||||||||||||||||||||||||||||||||||||||||||::||||||||||||
Db     781  CTGGAGTAATCCTACCACCACATGCCAACATAACTGAGATTCAATTACYGATGAAGGCA 840

Qy    4296  CTATCCCCTTCATGGAAAAGATTAAGGAGAAAATCTGAAGAAAGGGAGACACCTTA 4355
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841  CTATCCCCTTCATGGAAAAGATTAAGGAGAAAATCTGAAGAAAGGGAGACACCTTA 900

Qy    4356  TCTTTGAGGCTACCAAAAAACACTGTGATGAGGTTGCTAACGAGTTAGCTCGAAAGGAA 4415
            ||||||||||||||||||||||||||||||||||||||||||||||::||||||||||||
Db     901  TCTTTGAGGCTACCAAAAAACACTGTGATGAGGTTGCTAACGAGTTAGCTCGAAAGGGAA 960

Qy    4416  TAACAGCTGTCTCTTACTATAGGGATGTGACATCTCAAAATCCTGAGGGCGACTGTG 4475
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961  TAACAGCTGTCTCTTACTATAGGGATGTGACATCTCAAAAATCCTGAGGGCGACTGTG 1020

Qy    4476  TAGTAGTTGCCACTGACTTTGATCCGTGTATG 4535
            ||||||||||||||||||||||||||||||||
Db    1021  TAGTAGTTGCCACTGATCCCTTGTGTACAGGGTGACACTTTGATTCCGTGTATG 1080

Qy    4536  ACTGCAGGCTCATGGTGAGAAGGCACATGCCATGTTGACCCTACTTTCACCATGG 4595
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  1081 ACTGCCAGCCATGCGTAGAAGGCACATGCCATGTGACCTTGACCCTACTTTCACCATGG 1140

Qy  4596 GTGTTCGTGTGTGCGGGGTTTCAGCAATAGTTAAAGGCCAGCCTAGGGGCCGCACAGCC 4655
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 GTGTTCGTGTGTGCGGGGTTTCAGCAATAGTTAAAGGCCAGCCTAGGGGCCGCACAGCC 1200

Qy  4656 GTGGGAGAGCTGGCATATACTACTATGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTC 4715
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GTGGGAGAGCTGGCATATACTACTATGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTC 1260

Qy  4716 CTGAATGCAACATTGTTGAAGCCTTCGACGACGCCAAGCATGTATGGTTTGTCATCAA 4775
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 CTGAATGCAACATTGTTGAAGCCTTCGACGACGCCAAGCATGTATGGTTTGTCATCAA 1320

Qy  4776 CAGAAGCTCAAAACTATTCTGGACACCTATCGGCACCCAACCTGGGTTACCTGCGATAGGAG 4835
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CAGAAGCTCAAAACTATTCTGGACACCTATCGGCACCCAACCTGGGTTACCTGCGATAGGAG 1380

Qy  4836 CAAATTTGGACGAGTGGGCGATCTCTTTCTATGGTCAACCCGAACCTTCATTTGTCA 4895
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 CAAATTTGGACGAGTGGGCGATCTCTTTCTATGGTCAACCCGAACCTTCATTTGTCA 1440

Qy  4896 ATACTGCAAAAAGAAGAACTGCTGACAAATATGTTTGTTGACTGCAGCCCAACTACAACTGT 4955
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 ATACTGCAAAAAGAAGAACTGCTGACAAATATGTTTGTTGACTGCAGCCCAACTACAACTGT 1500

Qy  4956 GTCATCAGTATGGCTATGCTGCTCGCTCCCAATGACGCCACCACGGTGGCAGGGAGCCGGCCTG 5015
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 GTCATCAGTATGGCTATGCTGCTCGCTCCCAATGACGCCACCACGGTGGCAGGGAGCCGGCCTG 1560

Qy  5016 GGAAAAAACCTTGTGTGCGGGCGGTTCTGTGGGCGCTTGGACGGCGCTGGACGGCGCTGGCGCC 5075
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 GGAAAAAACCTTGTGTGCGGGCGGTTCTGTGGGCGCTTGGACGGCGCTGGACGGCGCTGGCGCC 1620

Qy  5076 AGCCAGCGAGGTGACCAGATACCAGATACCAGAGTGCTTCACTGAAGTCAATACTTCATTCCCTTGGACAG 5135
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 AGCCAGCGAGGTGACCAGATACCAGATACCAGAGTGCTTCACTGAAGTCAATACTTCATTCCCTTGGACAG 1680

Qy  5136 CCGGCACTCGCTGTTGGGCGTTGCGTTGGGCTTATCTAGCCCATTGACACTTTTGGCG 5195
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 CCGGCACTCGCTGTTGGGCGTTGCGTTGGGCTTATCTAGCCCATTGACACTTTTGGCG 1740

Qy  5196 CCAGAACTTGTTCAATTAACACCTTTCTTGGGCCTCATGGAGCTACTTGCTA 5255
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 CCAGAACTTGTTCAATTAACACCTTTCTTGGGCCTCATGGAGCTACTTGCTA 1800

Qy  5256 CAGTGGTTGACGAAGAAGAGAAAATCGTGAGGAGTGCATCAATTCATTCCCTTTGGAGCCA 5315
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 CAGTGGTTGACGAAGAAGAGAAAATCGTGAGGAGTGCATCAATTCATTCCCTTTGGAGCCA 1860

Qy  5316 TGGTTGCTGCAATTGACAAGCTGAAGAGTACAATCACCACAGCTAGTCCTTTCACCATTGG 5375
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 TGGTTGCTGCAATTGACAAGCTGAAGAGTACAATCACCACAGCTAGTCCTTTCACCATTGG 1920

Qy  5376 AAACCGCCCTTGAAAAAAACTTAACACCTTTCTTGGGCCTCATGGAGCTACAATCCTTGCTA 5435
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 AAACCGCCCTTGAAAAAAACTTAACACCTTTCTTGGGCCTCATGGAGCTACAATCCTTGCTA 1980

Qy  5436 TCATAGAGTATTGCTGTGGTTTAGTCACTTACCTTACCTGACAATCCCTTTGCATCATGCGTGT 5495
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 TCATAGAGTATTGCTGTGGTTTAGTCACTTACCTTACCTGACAATCCCTTTGCATCATGCGTGT 2040

Qy  5496 TTGCTTTCATTGCGGTATTCGGTCGTGCGTTATTACCCCCACTACCTCACAAGATCAAAATGTTCCTGTCAT 5555
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 TTGCTTTCATTGCGGTATTCGGTCGTGCGTTATTACCCCCACTACCTCACAAGATCAAAATGTTCCTGTCAT 2100

Qy  5556 TATTTGGAGGCGCAATTGCGTCCGCCAGGTTACGAGCGCACTGGCGTTCATGA 5615
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 TATTTGGAGGCGCAATTGCGTCCGCCAGGTTACGAGCGCCACTGGCGTTCATGA 2160

Qy  5616 TGGCCGGGCGCGGAACAGCTCTTGGTACATGCGGTACATCGGTTTGTCTTTTGACA 5675
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 TGGCCGGGCGCGYGGGAACAGCTCTTGGTACATCGGTACATCGGTTTGTCTTTTGACA 2220
```

```
Qy  5676 TGCTAGGGGCGCTATGCGCTGCCGCCTCATCCACTGCTTGCTTGACATTTAAATGCTTGATGG 5735
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 TGCTAGGGGCGCTATGCGCTGCCGCCTCATCCACTGCTTGCTTGACATTTAAATGCTTGATGG 2280

Qy  5736 GTGAGTGGCCACTATGATCAGATCAGCTTGCTGGTTAGTCTACTCCGCGTTCAATCCGGCCG 5795
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 GTGAGTGGCYCACTATGATCAGATCAGCTTGCTGGTTAGTCTACTCCGCGTTCAATCCGGCCG 2340

Qy  5796 CAGGAGTTGTGGGCGTCTTGTGTCAATGTTTGCTTGTGCAATGTTTGCTTTGACACAGCAGCAGGGCCAG 5855
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 CAGGAGTTGTGGGCGTCTTGTGTCAATGTTTGCTTGTGCAATGTTTGCTTTGACACAGCAGCAGGGCCAG 2400

Qy  5856 ATCACTGGCCACAGACTTCTTACTATGCTTGCTAGGAGCAACACTGTATGTAATGAGT 5915
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db  2401 ATCACTGGCCACAGACTTCTTACTATGCTTGCTAGGAGCAACACTGTATGTAARTGAGT 2460

Qy  5916 ACTTTATTGCCACTGTGACATCCGCAGGAAGATACTGGGCATTCTGGAGGCATCTACCC 5975
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 ACTTTATTGCCACTGTGACATCCGCAGGAAGATACTGGGCATTCTGGAGGCATCTACCC 2520

Qy  5976 CCTGGAGTGTCATATCAGCTTGCATCCGTTGGCTCCACACCCCGACGCGGAGGATGATTGCG 6035
         |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db  2521 CCTGGAGTRTCATATCAGCTTGCATCCGTTGGCTYCACACCCCGACGCGGAGGATGATTGCG 2580

Qy  6036 GCCTCATTGCTTGGGGTTAGAGATTTGGCAGTATGTGTCAATTTCTTTGTGTGATTTGCT 6095
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581 GCCTCATTGCTTGGGGTCTARAGATTTGGCAGTATGTGTCAATTTCTTTGTGTGATTTGCT 2640

Qy  6096 TTAATGTCCTTAAAGCTGAGTTCAGAGCATGGTTAACATTCCTGGTTGTGTCCTTTCTACA 6155
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641 TTAATGTCCTTAAAGCTGAGTTCAGAGCATGGTTAACATTCCTGGTTGTGTCCTTTCTACA 2700

Qy  6156 GCTCCAGAAGGGGTACAAGGGCCCCTGGATCAGTATGCTCCAAGCACGCGTC 6215
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 GCTCCAGAAGGGGTACAAGGGCCCCTGGATCAGTATGCTCCAAGCACGCGTC 2760

Qy  6216 CATGCGGTGCTGAACTCATCTTTCTGTTGAGAATGGTTTGCAAAACTTTACAAAGGAC 6275
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 CATGCGGTGCTGAACTCATCTTTCTGTTGAGAATGGTTTGCAAAACTTTACAAAGGAC 2820

Qy  6276 CCAGAACTTGTTCAAGTTACTGACGAGAGGCGCTGTTCCAGTCAACGCTGGCTGTGTGTGGGT 6335
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821 CCAGAACTTGTTCAAGTTACTGACGAGAGGCGCTGTTCCAGTCAACGCTGGCTGTGTGTGGGT 2880

Qy  6336 CGGCTAGACCGGACCGGACACTGATTGGACTAGTCTTGTCGTCAATTATGGCGGACT 6395
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881 CGGCTAGACCGGACCGGACACTGATTGGACTAGTCTTGTCGTCAATTATGGCGGACT 2940

Qy  6396 ACTGTAAATATGAGAAAATGGGAGATCACATTTTTGTTACAGCAGTATCCTCTCCAAATG 6455
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 ACTGTAAATATGAGAAAATGGGAGATCACATTTTTGTTACAGCAGTATCCTCTCCAAATG 3000

Qy  6456 TCTGTTTCACCCAGGTGCCCCAACCTTGAGAGCTGCGCAGTGGCCGTGACCGCGTACAGG 6515
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 TCTGTTTCACCCAGGTGCCCCAACCTTGAGAGCTGCGCAGTGGCCGTGACCGCGTACAGG 3060

Qy  6516 TTCAGTGTATCTAGGTGAGCCCAAAACTCCTTGCGTGTTACGGTC 6575
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 TTCAGYGTATCTAGGTGAGCCCAAAACTCCTTGCGTGTTACGGTC 3120

Qy  6576 CTGACGGTAAGGTAAAACTGTTAAGCTTCCCTTCCCTTCCGCGGTTGACGGTTACGGGTG 6635
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 CTGACGGTAAGGGTAAAACTGTTAAGCTTCCCTTCCCTTCCGCGGTTGACGGTTGCGGTG 3180

Qy  6636 TGGCCATGCAACTTAATTTGCGTGATGCAGCTTGAGGCAAATGACTGTAATTCCACAAACA 6695
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 GTCGCATGCAACTTAATTTGCGTGATGCAGCTTGAGGCAAATGACTGTAATTCCATAAACA 3240

Qy  6696 ACACTCCTAGTGATGATGAAGCCGCAGTGTCCGCTCTTGTTTTCAAACAGGAGTTGCGGCGTA 6755
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 ACACTCCTAGTGATGATGAAGCCGCAGTGTCCGCTCTTGTTTTCAAACAGGAGTTGCGGCGTA 3300
```

```
AC   AAT00045;
XX
DT   27-AUG-2003  (revised)
DT   02-JUL-1996  (first entry)
XX
DE   Hepatitis GB virus (HGBV) clone 13.
XX
KW   Hepatitis GB virus; HGBV; diagnosis; treatment; vaccine; reagents; non-A;
KW   non-B; non-C; non-D; non-E; clone 13; tamarin; infected plasma;
KW   lambda phage; cDNA library; ss.
XX
OS   Hepatitis G virus.
XX
PN   WO9521922-A2.
XX
PD   17-AUG-1995.
XX
PF   14-FEB-1995;   95WO-US002118.
XX
PR   14-FEB-1994;   94US-00196030.
PR   13-MAY-1994;   94US-00242654.
PR   29-JUL-1994;   94US-00283314.
PR   23-NOV-1994;   94US-00344185.
PR   23-NOV-1994;   94US-00344190.
PR   27-JAN-1995;   95US-00344557.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Simons JN, Pilot-Matias TJ, Dawson GJ, Schlauder GG, Desai SM;
PI   Leary TP, Muerhoff AS, Erker JC, Buijk SL, Mushahwar IK;
XX
DR   WPI; 1995-293123/38.
XX
PT   Non-A, non-B, non-C, non-D, non-E Hepatitis virus reagents - useful for
PT   diagnosis and therapy of hepatitis GB virus.
XX
PS   Example 5; Page 191; 661pp; English.
XX
CC   Double stranded hepatitis GB virus (HGBV) DNA obtd. from HGBV infected
CC   tamarin plasma, using standard procedures, was used to prepare a lambda
CC   phage HGBV cDNA library. The cDNA clone AAT00045, was rescued from the
CC   lambda phage, searched against a sequence database and found to be an
CC   unique HGBV sequence. Reagents which comprise the HGBV DNA, or its
CC   protein prods. can be used for the diagnosis, therapy or in a vaccine to
CC   prevent HGBV infection. (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 479 BP; 113 A; 133 C; 114 G; 118 T; 0 U; 1 Other;

Query Match             4.8%;  Score 449.8; DB 2; Length 479;
Best Local Similarity   97.5%;  Pred. No. 2.8e-121;
Matches  467;  Conservative   1;  Mismatches   8;  Indels   3;  Gaps   1;
```

```
Qy 6419 GATCACATTTTGTTACAGCAGTATCCTCCAAATGTCTGTTCACCCAGTGCCCCCA 6478
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GATCACATTTTGTTACAGCAGTATCCTCCAAATGTCTGTTCACCCAGTGCCCCCA 60

Qy 6479 ACCTTGAGAGCTGCGCCGTGGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCC 6538
        |||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db   61 ACCTTGAGAGCTGCGCCGTGGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCC 120

Qy 6539 AAAACTCCTTGACGACGCATCTGCTTGCTGTGTCCTGACGGTCCTGACGGTAAAACTGTT 6598
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AAAACTCCTTGACGACGCATCTGCTTGCTGTGTCCTGACGGTCCTGACGGTAAAACTGTT 180

Qy 6599 AAGCTTCCCTTCCGCGTTGACGGTGTGCACACACCTGGTGCTGTGGCATGCAACTTAATTGGGT 6658
        ||||||||||||||||||||||||||||||                |||||||||||||||||
Db  181 AAGCTTCCCTTCCGCGTTGACGGTGTGCATGCAACTTAATTGGGT 240

Qy 6659 AAGCTTCCCTTCCGCGTTGACGGTGTCACACACCTGGTGCTGTGGCATGCAACTTAATTGGGT 6718
        ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db  241 AAGCTTCCCTTCCGCGTTGACGGTGTAATTCCATAAACAACACTCCTAGTGATGAAGCCGCA 300

Qy 6719 GTGTCCGCTCTTGTTGTTTCAAACGAGTTGCGCGGCTACAAACCAATTGCTTGAGGCAATT 6778
```

```
Qy 6756 CAAACCAATTGCTTGAGGCAATTTCAGCTGCCGTTGACACCACCAAACTGCCAGCCCCCT 6815
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3301 CAAACCAATTGCTTGAGGCAATTTCAGCTGCCGTTGACACCACCAAACTGCCAGCCCCCT 3360

Qy 6816 CC---ATCGAAGAGGTAGTGTGTAAGAAAGCGCCAGTTCCGGGCAAGAACTGGTTCGCTTA 6872
        ||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3361 CCCAGATCGAAGAGGTAGTGTGTAAGAAAGCGCCAGTTCCGGGCAAGAACTGGTTCGCTTA 3420

Qy 6873 CCTTGCCTCCCCCTCCGAGATCCGTCCAGAGAGTCGTCATGTCCTGAAAGCTGCAACGAA 6932
        ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db 3421 CCTTGCCTCCCCCTCCGAGATCCGTCCAGAGAGTGTCATGTCCTGAAAGCCTGCAACGAA 3480

Qy 6933 GTGACCCGTTAGAAGGTCCTTCAAACCTCCCTCCTTCACCACCTGTTCTACAGTTGGCCA 6992
        ||||:||||||||||||||:|||||||||||||  :|||||||||||||::|||||||||
Db 3481 GTGAATGGTCAGACGAAAGTTGGTCAACCTCCCTCTGTTCTRCAGTTGGCCA 3540

Qy 6993 TGCCGATGCCCCTGTTGGGAGCCGGGTGAGTGTAAACCCTTTCACTGACGCAATTGGATGTGCAA 7052
        ||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||||||
Db 3541 TGCCGATGCCCCTGTTGGGAGCAGTGAGTGTAAACCCTTTCACTGACGCAATTGGATGTGCAA 3600

Qy 7053 TGACCGAAACAGGGCGGAGGCCCTGATGATTTACCCGTTACCCTCCCAAAAAGGAGGTCT 7112
        |||||||||:||:  ||||||||| : :|||||||||:||||||||||||||||||||||
Db 3601 TGACCGAAACARGYGGAGKCCCWSAKRATTTACCCAGTTACCCTCCCAAAAAGGAGGTCT 3660

Qy 7113 CTGAATGGTCAGACGAAAGTTGGTCGACGGCTACAACCGCTTCCAGCTACGTTACTGGCC 7172
        |||||||||||||||||||||||||:||:|||||||||||||||||||||||||||||||
Db 3661 CTGAATGGTCAGACGAAAGTTGGTCAACGACTACAACCGCTTCCAGCTACGTTACTGGCC 3720

Qy 7173 CCCCGTACCCTAAGATACGGGAAAGGATTCCACTCAGTCAGCCCCGCCAAACGGCCTA 7232
        ||||||||||||||||||||| |||||||||||||| ||||||||||||||||||||||
Db 3721 CCCCGTACCCTAAGATACGGGCAAGGATTCCACTCAATCAGCCCAAACGGCCTA 3780

Qy 7233 CAAAAAGAAGTGGGAAAAGATGGAGTGAGTTTCGTGCAGCATGAGCTACACCTGACCGACG 7292
        |||||||||:||||||||:|||||||||||||||||||||||||||||||||:|||||||||
Db 3781 CAAAAAGAAGAAGTGGGAAAGATGGAGTGAGTTTCGTGCAGCATGAGCTACACTTGGACCGACG 3840

Qy 7293 TGATTAGCTTCAAAACTGCTTCTAAAGTTCGTGTCGCAACTCGGGCCATCACTAGTGGTT 7352
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3841 TGATTAGCTTCAAAACTGCTTCTAAAGTTCGTGTCGCAACTCGGGCCATCACTAGTGGTT 3900

Qy 7353 TCCTCAAACAAAGATCATTGGTACTGACTGAGCGCCGCGGGGATGCGGAGCTTAGAAAAC 7412
        |||||||||||||||||||||| ||||||||||||||| |||||||||||||||||||||
Db 3901 TCCTCAAACAAAGATCATTGGTGTATGTGACTGAGCGCCGGGATGCGGAGCTTAGAAAAC 3960

Qy 7413 AAAAAGTCACTATTATTAGACACACCCTCGTTCCCCCCCATCATCACCACAAGCAAGTGAGAT 7472
        ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db 3961 AAAAAGTCACTATTATTAGACAACACCTCGTTCCCCCCCATCATCACCACAAGCAAGTGAGAT 4020

Qy 7473 TGGCTAAGGAAAAAGCTTCAAAAGTTCGGTGTCATGTGGGACTATGGAGTAGTAGCAG 7532
        ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db 4021 TGGCTAAGGAAAAAGCTTCAAAAGTTCGGTGTCATGTGGGACTATGATGGAAGTAGTAGCAG 4080

Qy 7533 CTCACACGCCCTCTAGTCTGCTAAGTCCCACATCACTGGGCCTTCGGGGCACTGATGTTC 7592
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4081 CTCACACGCCCTCTAGTCTGCTAAGTCCCACATCACTGGGCCTTCGGGGCACTGA----- 4135

Qy 7593 GTTCTGGAGCAGCAGCCCCGCCCAAGGCTGTTCTGCAGAAGTGTGTCGAGGCAGGTGAGA 7652
                        ||||||||||||||||||||||||||||||
Db 4136 ----------------TGTTCTGGACTTGCAGAAGTGTGTCGAGGCAGGTGAGA 4173

Qy 7653 TACCGAGTCATTATCGGCAAACTGTGATAGTTCCAAAGTGAGGAGGTCTTCGTGAAGACCC 7712
        ||||||||||||||||||||||||||||||||||||||           |||||||||||
Db 4174 TACCGAGTCATTATCGGCAAACTGTGATAGTTCCAATAAACAACACTCCGTGAAGACCC 4233

Qy 7713 CCCAGAAACCAACAAAGAAACCCCCAAGGCTTATC 7747
        |||||||||||||||||||||||||||||||||||
Db 4234 CCCAGAAACCAACAAAGAAACCCCCAAGGCTTATC 4268
```

```
RESULT 15
AAT00045
ID   AAT00045  standard; DNA; 479 BP.
XX
```

```
Db        301 GTGTCCGCTCTTGTTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATT 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy       6779 TCAGCTGGCGTTGACACCACCAAACTGCCAGCCCCTCC---ATCGAAGAGGTAGTCGTA 6835
              ||||||||||||||||||||||||||||||||||||    |||||||||||||||||

Db        361 TCAGCTGGCGTTGACACCACCAAACTGCCAGCCCCTCCCCAGATCGAAGAGGTAGTCGTA 420

Qy       6836 AGAAAGCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCTCCCGAGATC 6894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        421 AGAAAGCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCTCCCGAGATC 479
```

Search completed: October 30, 2005, 11:07:26
Job time : 2963 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:       October 30, 2005, 08:56:10 ; Search time 25672 Seconds
                                           (without alignments)
                                           17740.362 Million cell updates/sec

Title:        US-10-009-002-1
Perfect score: 9399
Sequence:     1 accacaaacactccagttg........cccgcttggaattaaaaact 9399

Scoring table: IDENTITY NUC
               Gapop 10.0 , Gapext 1.0

Searched:     4708233 seqs, 2422760755 residues

Total number of hits satisfying chosen parameters:      9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    GenEmbl:*
              1: gb_ba:*
              2: gb_htg:*
              3: gb_in:*
              4: gb_om:*
              5: gb_ov:*
              6: gb_pat:*
              7: gb_ph:*
              8: gb_pl:*
              9: gb_pr:*
             10: gb_ro:*
             11: gb_sts:*
             12: gb_sy:*
             13: gb_un:*
             14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	%Query Match	Length	DB	ID	Description
1	9399	100.0	9399	6	AX057392	AX057392 Sequence
2	9399	100.0	9399	12	AF179612	AF179612 Hepatitis
3	9376.6	99.8	9399	12	AY243572	AY243572 Synthetic
4	9338.6	99.4	9397	6	AX805212	AX805212 Sequence
5	9338.6	99.4	9397	14	HGB277947	AJ277947 Hepatitis
6	9133	97.2	9139	6	AX057394	AX057394 Sequence
7	9059.8	96.4	9143	6	AR075140	AR075140 Sequence
8	9059.8	96.4	9143	6	AR230431	AR230431 Sequence
9	9059.8	96.4	9143	6	AR230434	AR230434 Sequence
10	9059.8	96.4	9143	6	AR310126	AR310126 Sequence
11	9059.8	96.4	9143	6	AR310129	AR310129 Sequence
12	9059.8	96.4	9143	6	AR350538	AR350538 Sequence
13	9059.8	96.4	9143	6	AR350541	AR350541 Sequence
14	9059.8	96.4	9143	6	AR494272	AR494272 Sequence
15	9059.8	96.4	9143	6	AR494275	AR494275 Sequence
16	9059.8	96.4	9143	14	HGU22304	U22304 Hepatitis G
17	8692.6	92.5	8912	6	AR230335	AR230335 Sequence
18	8692.6	92.5	8912	6	AR310030	AR310030 Sequence
19	8692.6	92.5	8912	6	AR350442	AR350442 Sequence
20	8692.6	92.5	8912	6	AR494176	AR494176 Sequence
21	8027	64.8	8027	14	GVI428955	AJ428955 Hepatitis
22	6088	64.8	8069	6	AX805211	AX805211 Sequence
23	6088	44.2	4268	6	AR230353	AR230353 Sequence
24	4159	44.2	4268	6	AR310048	AR310048 Sequence
25	4159	44.2	4268	6	AR350460	AR350460 Sequence
26	4159	44.2	4268	6	AR494194	AR494194 Sequence
27	1223.4	13.0	1233	14	AY534875	AY534875 Hepatitis
28	1223.4	13.0	1233	14	AY534877	AY534877 Hepatitis
29	1217	12.9	1233	14	AY534878	AY534878 Hepatitis
30	1215.4	12.9	1233	14	AY534873	AY534873 Hepatitis
31	1215.4	12.9	1233	14	AY534874	AY534874 Hepatitis
32	1215.4	12.9	1233	14	AY534876	AY534876 Hepatitis
33	449.8	4.8	479	6	AR230346	AR230346 Sequence
34	449.8	4.8	479	6	AR310041	AR310041 Sequence
35	449.8	4.8	479	6	AR310128	AR310128 Sequence
36	449.8	4.8	479	6	AR350453	AR350453 Sequence
37	449.8	4.8	479	6	AR350540	AR350540 Sequence
38	449.8	4.8	479	6	AR494187	AR494187 Sequence
39	449.8	4.8	479	6	AR494274	AR494274 Sequence
40	449.8	4.8	479	14	Y18973	Y18973 Hepatitis G
41	343.4	3.7	357	14	HGB18973	HGB18973
42 c	319.6	3.4	337	6	AR230350	AR230350 Sequence
43 c	319.6	3.4	337	6	AR230426	AR230426 Sequence
44 c	319.6	3.4	337	6	AR310045	AR310045 Sequence
45 c	319.6	3.4	337	6	AR310121	AR310121 Sequence

ALIGNMENTS

```
RESULT 1
AX057392
LOCUS       AX057392                9399 bp    DNA     linear   PAT 17-JAN-2001
DEFINITION  Sequence 1 from Patent WO0075337.
ACCESSION   AX057392
VERSION     AX057392.1  GI:12310132
KEYWORDS
SOURCE      Hepatitis GB virus B
  ORGANISM  Hepatitis GB virus B
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae.
REFERENCE   1
  AUTHORS   Bukh,J., Yanagi,M., Emerson,S.U. and Purcell,R.H.
  TITLE     Infectious cdna clone of gb virus b and uses thereof
  JOURNAL   Patent: WO 0075337-A 1 14-DEC-2000;
            THE GOVERNMENT OF THE UNITED STATES OF AMERICA (US)
FEATURES             Location/Qualifiers
     source          1..9399
                     /organism="Hepatitis GB virus B"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:39113"
ORIGIN

Query Match       100.0%;  Score 9399;  DB 6;  Length 9399;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 9399;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG   60

Qy   61 CAGGGCGTGGGGGATTCCCCTGCCCGTCTGCAGAAGAGGGTGAGCCAACCACCTTAGTAT  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CAGGGCGTGGGGGATTCCCCTGCCCGTCTGCAGAAGAGGGTGAGCCAACCACCTTAGTAT  120

Qy  121 GTAGGCCGCGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTAGGCCGCGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180

Qy  181 CCTGATGGGCGTTCATGGGTTCGGTGGTGTGGTGCGCTTTAGGCAGCCTCCACGCCCACCA  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CCTGATGGGCGTTCATGGGTTCGGTGGTGTGGTGCGCTTTAGGCAGCCTCCACGCCCACCA  240
```

```
Qy  1321 CGAGGCTGTCATCTTCTTGACCAAACTGCTTCACAAGTACCATACGCTATTGCGACTAT 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CGAGGCTGTCATCTTCTTGACCAAACTGCTTCACAAGTACCATACGCTATTGCGACTAT 1380

Qy  1381 GTTTAGCAGTGTACACTACCTGGCGGTTGGCGCTGGCGGTTGGCGCTGGTCGGGGGCAA 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 GTTTAGCAGTGTACACTACCTGGCGGTTGGCGCTGGCGGTTGGCGCTGGTCGGGGGCAA 1440

Qy  1441 GTGTATCAGTTGCTCCTAGCGCTTACATAGAAGCGACCTCTGGAAACCCCAT 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 GTGTATCAGTTGCTCCTAGCGCTTACATAGAAGCGACCTCTGGAAACCCCAT 1500

Qy  1501 CAGGGTGCCCACTGGATGCTCAATAGCGTCGAGTTTGCTGCGCCTTGATGATACCATGTCC 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 CAGGGTGCCCACTGGATGCTCAATAGCGTCGAGTTTGCTGCGCCTTGATGATACCATGTCC 1560

Qy  1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTGTACAGTGCCAAAGTGGAC 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTGTACAGTGCCAAAGTGGAC 1620

Qy  1621 CAGGCCTATCACTCTAGATATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 CAGGCCTATCACTCTAGATATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG 1680

Qy  1681 TGGCGAGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 TGGCGAGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA 1740

Qy  1741 TGTGCCATCGTACTGCACTATGGCACTGATGCGCAGTGTGGAACGACACTCGCAACACTTA 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 TGTGCCATCGTACTGCACTATGGCACTGATGCGCAGTGTGGAACGACACTCGCAACACTTA 1800

Qy  1801 CGAAGCATGCGGTGTAACACCATGGCTAACACCGGCACAACGGCTCAGCCTGAA 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 CGAAGCATGCGGTGTAACACCATGGCTAACACCGGCACAACGGCTCAGCCTGAA 1860

Qy  1861 ATTGGCTATATTACAATACCCTGGGTCTCAAAGAAAATGTTAAACCTCATAATTGGATGTC 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 ATTGGCTATATTACAATACCCTGGGTCTCAAAGAAAATGTTAAACCTCATAATTGGATGTC 1920

Qy  1921 AGGCCATTGTATTTGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAA 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 AGGCCATTGTATTTGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAA 1980

Qy  1981 TTTCACTCTCCTACCGGAGAGGTGGCTAGGTTGCCGGTACCCCACCTGTGGTACG 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 TTTCACTCTCCTACCGGAGAGGTGGCTAGGTTGCCGGTACCCCACCTGTGGTACG 2040

Qy  2041 TGGTTCTTGGTTACAGGTTCCGCAAGAGGTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 TGGTTCTTGGTTACAGGTTCCGCAAGAGGTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100

Qy  2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGTTCTTATATTCCGCCACGGGTGC 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGTTCTTATATTCCGCCACGGGTGC 2160

Qy  2161 TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGTGCTAATTCGTGTTGTGTGG 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGTGCTAATTCGTGTTGTGTGG 2220

Qy  2221 CAGCAAGTATCTTATTTTAGCCTTGGTGTGTATTAGTCGGTGACTG 2280
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 CAGCAAGTATCTTATTTTAGCCTTGGTGTGTATTAGTCGGTGACTG 2280

Qy  2281 TGGTTACCCTTGGCTCCGTGCTCCATCCCAAGCTGGCTCGGATGT 2340
         |||||||||||||||||||||||||||||||||||||||||||||
Db  2281 TGGTTACCCTTGGCTCCGTGCTCCATCCCAAGCTGGCTCGGATGT 2340

Qy  2341 TTTGTCTAAAGCTCAAGTAGCTCCCTTTTGATTTTGATTTGTCTATCTCCG 2400
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 TTTGTCTAAAGCTCAAGTAGCTCCCTTTTGATTTTGATTTGTCTATCTCCG 2400

Qy  2401 CTGCAGGCTACGTTATGCGTGTGTGTGCCCATGGGTTGTGCCGCAGGCTGCCCCT 2460
```

```
Qy   241 CCTCCCGATAGAGCGGCGGCACTGTAGGGAAGACCGGGGACCGGTCACTACCAAGGACG 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCCCGATAGAGCGGCGGCACTGTAGGGAAGACCGGGGACCGGTCACTACCAAGGACG 300

Qy   301 CAGACCTCTTTTGAGTATCACGCCTCCGGAGTAGTTGGGCAAGCCCACCTATATGTGT 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGACCCCTTTTGAGTATCACGCCTCCGGAGTAGTTGGGCAAGCCCACCTATATGTGT 360

Qy   361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACCGTACCATCACCGTACCGTACTGCCTGA 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACCGTACCATCACCGTACCGTACTGCCTGA 420

Qy   421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACC 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACC 480

Qy   481 TGGCCCCAGAACGCGCAAGAACAAGCAGACGCAGGCTTCATATCCTGTCCATTAAAAC 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TGCGCCCCAGAACGCGCAAGAACAAGCAGACGCAGGCTTCATATCCTGTCCATTAAAAC 540

Qy   541 ATCTGTTGAAAGGGGACAACGCAAAGCCAGCGGAGGCGATGCTCGGCCTCGTAA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 ATCTGTTGAAAGGGGACAACGCAAAGCCAGCGGAGGCGATGCTCGGCCTCGTAA 600

Qy   601 TTACAAAATTGCTGTATCCGACATTGGCTCAGCTGCTTTGCCAGC 660
         |||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTACAAAATTGCTGTATCCGACATTGGCTCAGCTGCTTTGCCAGC 660

Qy   661 TCATGGTTGGGGACGCCAAGACCCTCGCAATCTTGGAATCTTCTGA 720
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TCATGGTTGGGGACGCCAAGACCCTCGCAATCTTGGAATCTTCTGA 720

Qy   721 TTACCCTTGGGGTGGATTGGTGATGTTACAACTCACACCTCTCAGTAGGGCCCGCTGGT 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TTACCCTTGGGGTGGATTGGTGATGTTACAACTCACACCTCTCAGTAGGGCCCGCTGGT 780

Qy   781 GGCAGGAGCGGTCGTTCGACCAGTCTGCCAGATAGTACGCTTGCTGGAGGATGGAGTCAA 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GGCAGGAGCGGTCGTTCGACCAGTCTGCCAGATAGTACGCTTGCTGGAGGATGGAGTCAA 840

Qy   841 CTGGGCTACTGGTTCGGTGTCCACCTTTTGTGTGGTATGTCTGCTATCTTCTTGGCCTG 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CTGGGCTACTGGTTCGGTGTCCACCTTTTGTGTGGTATGTCTGCTATCTTCTTGGCCTG 900

Qy   901 TCCCTGTAGTGGGCGCGGGGTCACTGACCGACCAGAACAAATACCACAATCCTGACCAATTG 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCCTGTAGTGGGCGCGGGGTCACTGACCGACCAGAACAAATACCACAATCCTGACCAATTG 960

Qy   961 CTGCCAGCGTAATCAGTTTATCTATTGGTGAGTTGCGTGTTATTAGTCGGTGACTG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CTGCCAGCGTAATCAGTTTATCTATTGGTGAGTTGCGTGTTATTAGTCGGTGACTG 1020

Qy  1021 TGTGATCTGTGCGGACGAGTGCTGTGGGGTCAATCGGTACATCGGTACTTGTTACCTTGA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 TGTGATCTGTGCGGACGAGTGCTGTGGGGTCAATCGGTACATCGGTACTTGTTACCTTGA 1080

Qy  1081 TTTGGACTGGCACGGACACGGACTCCTTGGCTGACCCACCATTGATTTGTTATGGGGGTCTGT 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TTTGGACTGGCACGGACACGGACTCCTTGGCTGACCCACCATTGATTTGTTATGGGGGCTGT 1140

Qy  1141 GACCTGTGACGCGCCCTTGACATTGGTGAGTTGCGTGTATTAGTCGGTGACTG 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 GACCTGTGACGCGCCCTTGACATTGGTGAGTTGCGTGTATTAGTCGGTGACTG 1200

Qy  1201 GCTTGTCATGCAGGGCACGTTGCTTATTCAATGAAACTGGTACTTGTTACCTTGA 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GCTTGTCATGCAGGGCACGTTGCTTATTCAATGAAACTGGTACTTGTTACCTTGA 1260

Qy  1261 AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT 1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT 1320

Qy  1321 CGAGGCTGTCATCTTCTTGACCAAACTGCTTCACAAGTACCATACGCTATTGCGACTAT 1380
```

```
Qy 3541 GTCCCTTACTCCGTGCTCTTGCGGGAGACCAAGGGTATCTGGTAACGACGTGGGGTC 3600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3541 GTCCCTTACTCCGTGCTCTTGCGGGAGACCAAGGGTATCTGGTAACGACGTGGGGTC 3600

Qy 3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGCGGGCCCTTCCCAT 3660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGCGGGCCCTTCCCAT 3660

Qy 3661 GGCTGTTGCCAAGGGTCTTCAGGTGCCCGATTCTGTGCTCTGTGCCTGCTTCAGC 3720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3661 GGCTGTTGCCAAGGGTCTTCAGGTGCCCGATTCTGTGCTCTGTGCCTGCTTCAGC 3720

Qy 3721 GATGTTCACCGCTGCTAGAAATTCTGGCGGTTCAGTCAGTCAGATTAGGGTTAGGCCGTT 3780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3721 GATGTTCACCGCTGCTAGAAATTCTGGCGGTTCAGTCAGTCAGATTAGGGTTAGGCCGTT 3780

Qy 3781 GGTGTGTCTGGATACCATCCCCAGTACACAGCACATGGCCACTCTTGATACAAAACCTAC 3840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3781 GGTGTGTCTGGATACCATCCCCAGTACACAGCACATGGCCACTCTTGATACAAAACCTAC 3840

Qy 3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC 3900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC 3900

Qy 3901 CAAATTACCACTTTCTTACATGCAGGAGAAGTAGGTCTTGGTCCTCTAAATCCCAGTGT 3960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3901 CAAATTACCACTTTCTTACATGCAGGAGAAGTAGGTCTTGGTCCTCTAAATCCCAGTGT 3960

Qy 3961 GGCTACAACAGATCAATGCCAAAGTACATGCGACGCGGCTACGGCGTGAATCAAATTG 4020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3961 GGCTACAACAGATCAATGCCAAAGTACATGCGACGCGGCTACGGCGTGAATCAAATTG 4020

Qy 4021 CTATTTTAATGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGGCAT 4080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4021 CTATTTTAATGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGGCAT 4080

Qy 4081 GTACCTGACCGGAGCATGTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGC 4140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4081 GTACCTGACCGGAGCATGTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGC 4140

Qy 4141 TACCGATGCAACCACCGTGTTGGGCATTGGAAAGGTCCATCCAAAAA 4200
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 4141 TACCGATGCAACCACCGTGTTGGGCATTGGAAAGGTCCATCCAAAAA 4200

Qy 4201 TGTTAGGCTAGTGGTTCTTGCCACGCGCTACCCCCCGGAGTAATCCCTACACCACATGC 4260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4201 TGTTAGGCTAGTGGTTCTTGCCACGCGCTACCCCCCGGAGTAATCCCTACACCACATGC 4260

Qy 4261 CAACATAACTGACATCTCAATTCAATTAACCGATGAAGGGAGCACCTTATCCCTTTCATGGGAAAAAGAT 4320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4261 CAACATAACTGACATCTCAATTCAATTAACCGATGAAGGGAGCACCTTATCCCTTTCATGGGAAAAAGAT 4320

Qy 4321 TAAGGAGGAAAATCTGAAGAAAGGAGCACACCTTATCTTTGAGGCTACCAAAAAAACACTG 4380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4321 TAAGGAGGAAAATCTGAAGAAAGGAGCACACCTTATCTTTGAGGCTACCAAAAAAACACTG 4380

Qy 4381 TGATGAGCTTGCTGACTAACGAGTTAGCTGAAAACAGCTGTCTCTTCTTACTATAGGGG 4440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4381 TGATGAGCTTGCTGACTAACGAGTTAGCTGAAAACAGCTGTCTCTTCTTACTATAGGGG 4440

Qy 4441 ATGTGACATCTCAAAAATCCCTAGAGTAGTAGTTGCCACTGAGTTGCCACTGCCTTGTG 4500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4441 ATGTGACATCTCAAAAATCCCTAGAGTAGTAGTTGCCACTGAGTTGCCACTGCCTTGTG 4500

Qy 4501 TACAGGGTACACTGGTACTTTGATTCCGTCATGGTAGAAGGCAC 4560
         ||||||||||||||||||||||||||||||||||||||||||||
Db 4501 TACAGGGTACACTGGTACTTTGATTCCGTCATGGTAGAAGGCAC 4560

Qy 4561 ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTGTTCGTGTGCGGGGTTTTCAGC 4620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4561 ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTGTTCGTGTGCGGGGTTTTCAGC 4620
```

```
Db 2401 CTGCAGGCTACGTTATGCTGCCCTTTAGGGTTGTGCCCATGGCTGCGGGCTTGCCCCT 2460

Qy 2461 AACTTTCTTGTTGCAGCAGCTGCTGCCCAACCAGATTATGACTGTGGGTGCGACTGCT 2520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2461 AACTTTCTTGTTGCAGCAGCTGCTGCCCAACCAGATTATGACTGTGGGTGCGACTGCT 2520

Qy 2521 AGTGGCAGGGTTAGTTTGTTTGTGGGCCGGCCGTAACCGGTCACCGGATAGCTCTGCTTGT 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2521 AGTGGCAGGGTTAGTTTGTTTGTGGGCCGGCCGTAACCGGTCACCGGATAGCTCTGCTTGT 2580

Qy 2581 AGTCCTTGGCCTCTGGTAGCGCTTTAACCCTTCTTGCATTTGGTTACGCGGCCTGCTTCAGC 2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2581 AGTCCTTGGCCTCTGGTAGCGCTTTAACCCTTCTTGCATTTGGTTACGCGGCCTGCTTCAGC 2640

Qy 2641 TTTTGATACCGAGATAATTGAGGGGCTGACATAACACCTGTAGTAGCATTAGTTGTCAT 2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2641 TTTTGATACCGAGATAATTGAGGGGCTGACATAACACCTGTAGTAGCATTAGTTGTCAT 2700

Qy 2701 GTCTCGTTTGCTTCTTTGCTCACTCGCTGCTTAGTTAGTTAACTCCTATCT 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2701 GTCTCGTTTGCTTCTTTGCTCACTCGCTGCTTAGTTAGTTAACTCCTATCT 2760

Qy 2761 TTGGCAACGTTGGGAGGAATTGGTTTGGGGAACGTTACACTAAGACCGGAGAGGTTTTTCCT 2820
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761 TTGGCAACGTTGGGAGGAATTGGTTTGGGGAACGTTACACTAAGACCGGAGAGGTTTTTCCT 2820

Qy 2821 TGTGCGGTTGTTTCCCCGGTGCGACATATGACCGCGCTGTGACTTTCTGTGTGTGTCA 2880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2821 TGTGCGGTTGTTTCCCCGGTGCGACATATGACCGCGCTGTGACTTTCTGTGTGTGTCA 2880

Qy 2881 CGTAGCTCTTCTATGTTTAACATCCAGTGCAGCATCGTTCTTTGGACTGACTCTAGGGT 2940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2881 CGTAGCTCTTCTATGTTTAACATCCAGTGCAGCATCGTTCTTTGGACTGACTCTAGGGT 2940

Qy 2941 TAGGGCCCATAGAATGTTGGTGCGTCTCGGAAAGTGTCATGCTTGGTATCCTCATTATGT 3000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2941 TAGGGCCCATAGAATGTTGGTGCGTCTCGGAAAGTGTCATGCTTGGTATCCTCATTATGT 3000

Qy 3001 TCTTAAGTTTTTCCTCTTAGTGTTTTGGTGAGAATGTGTGTTTTTCTATAAGCACTGCA 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3001 TCTTAAGTTTTTCCTCTTAGTGTTTTGGTGAGAATGTGTGTTTTTCTATAAGCACTGCA 3060

Qy 3061 TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCCATTTTTCCC 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3061 TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCCATTTTTCCC 3120

Qy 3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGGAATGGAATGAAGAACGCTTGGCGTGTGGGACAC 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGGAATGGAATGAAGAACGCTTGGCGTGTGGGACAC 3180

Qy 3181 GGTTGATGGTTGCCCGTGTTGTGCCGGCTCTCGGGACCTTGTTTCGCAGGGTTGGCTAT 3240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3181 GGTTGATGGTTGCCCGTGTTGTGCCGGCTCTCGGGACCTTGTTTCGCAGGGTTGGCTAT 3240

Qy 3241 GCCGCCAGATGGGTGGGCCATTACCGGCAACGTGCCAGTGTCTCTCTGAACGTGG 3300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3241 GCCGCCAGATGGGTGGGCCATTACCGGCAACGTGCCAGTGTCTCTCTGAACGTGG 3300

Qy 3301 CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC 3360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3301 CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC 3360

Qy 3361 TATCTTCAGATAGGATCTCTCGGCGTTGTGCGCGCTGTTTGTTGTGTGTGTGTCAAACGTGTT 3420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3361 TATCTTCAGATTAGGATCTCTCGGCGTTGTGCGCGCTACTAGCTAACGTGTTGTGTGTCAAACGTGTT 3420

Qy 3421 GTATACTGCTCACCATGGCGCGGCTAATGACCAAGGAGACATCTATCAACCAACCAGTGG 3480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3421 GTATACTGCTCACCATGGCGCGGCTAATGACCAAGGAGACATCTATCAACCAACCAGTGG 3480

Qy 3481 CCCAATAACCGTTGACGCGGCTAATGACCCAGGAGACATCTATCAACCAACCATGGGCTGG 3540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3481 CCCAATAACCGTTGACGCGGCTAATGACCCAGGAGACATCTATCAACCAACCATGGGCTGG 3540
```

```
Qy  4621  AATAGTTAAAGGCCAGCGTAGGGGCCCACAGGCCGTGGGAGAGCTGCATATACTACTA  4680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621  AATAGTTAAAGGCCAGCGTAGGGGCCCACAGGCCGTGGGAGAGCTGCATATACTACTA  4680

Qy  4681  TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT  4740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681  TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT  4740

Qy  4741  CGACGCAGCCAAGGCATGGTATGGTTTGTCATCAACAGAAGCTCAAACTATTCTGGACAC  4800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741  CGACGCAGCCAAGGCATGGTATGGTTTGTCATCAACAGAAGCTCAAACTATTCTGGACAC  4800

Qy  4801  CTATCGGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGCTGATCT  4860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801  CTATCGGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGCTGATCT  4860

Qy  4861  CTTTTCTATGGTCAACCCGAACCTTCATTGTCAATACTGCAAAAAGAACTGCTGACAA  4920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861  CTTTTCTATGGTCAACCCGAACCTTCATTGTCAATACTGCAAAAAGAACTGCTGACAA  4920

Qy  4921  TTATGTTGTTGACTGCAGCCCAACTACAACTGTGTCATCAGTAGTCATGCGTGCTCC  4980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921  TTATGTTGTTGACTGCAGCCCAACTACAACTGTGTCATCAGTAGTCATGCGTGCTCC  4980

Qy  4981  CAATGACGCCACCGGTGCGCAGGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG  5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981  CAATGACGCCACCGGTGCGCAGGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG  5040

Qy  5041  GCGCTTGACGGCGCTGTCCTGGCCCTGTGCGCGAGCCCAGCGGTGACCAGATACCA  5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041  GCGCTTGACGGCGCTGTCCTGGCCCTGTGCGCGAGCCCAGCGGTGACCAGATACCA  5100

Qy  5101  AATGTGCTTCACTGAAGTCAATACTTCTCGGACAGCAGCCGTCTCGCCGGTTGGGAGT  5160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101  AATGTGCTTCACTGAAGTCAATACTTCTCGGACAGCAGCCGTCTCGCCGGTTGGGAGT  5160

Qy  5161  GGCTATGGCTTATCTAGCACTTTGACACTTTGGCGCCCACTTGTGCGCGTTGCTGCTGTC  5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161  GGCTATGGCTTATCTAGCACTTTGACACTTTGGCGCCCACTTGTGCGCGTTGCTGCTGTC  5220

Qy  5221  TATTACATCAGTCCTACCGGTGCTACTGTCGCCCCCAGTGGTTGACGAAGAGAAATCGT  5280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221  TATTACATCAGTCCTACCGGTGCTACTGTCGCCCCCAGTGGTTGACGAAGAGAAATCGT  5280

Qy  5281  GGAGGAGTGTGCATCAATCCCTTGGAGGCCATGGTTGCTGCAATTGACAAGCTGAA  5340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281  GGAGGAGTGTGCATCAATCCCTTGGAGGCCATGGTTGCTGCAATTGACAAGCTGAA  5340

Qy  5341  GAGTACAATCACCACAACTAGTCCTTTCACATTGGAAACCGGCCAATTGCGTCCAA  5400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341  GAGTACAATCACCACAACTAGTCCTTTCACATTGGAAACCGGCCAATTGCGTCCAA  5400

Qy  5401  CTTTCTTGGGCCTCATGCAGCTACAACATCCTTGCTATCATAGAGTATTGCTGTGGTTAGT  5460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401  CTTTCTTGGGCCTCATGCAGCTACAACATCCTTGCTATCATAGAGTATTGCTGTGGTTAGT  5460

Qy  5461  CACTTTACCTGACAATGCAATCCCTTGCATCATGCGTTTCATTGGATGCCGGGTATTACTAC  5520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461  CACTTTACCTGACAATGCAATCCCTTGCATCATGCGTTTCATTGGATGCCGGGTATTACTAC  5520

Qy  5521  CCCACTACCTCACAAGATCAAGATCAAAATGTTCCTGTCTATTATTTGGAGGCGCAATTGCGTCCAA  5580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521  CCCACTACCTCACAAGATCAAGATCAAAATGTTCCTGTCTATTATTTGGAGGCGCAATTGCGTCCAA  5580

Qy  5581  GCTTACAGACGCTAGAGCGCCACTGCGTCGGCGTTCAGTGATGGCCGGGCTCGCGGAACAGCTCT  5640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581  GCTTACAGACGCTAGAGCGCCACTGCGTCGGCGTTCAGTGATGGCCGGGCTCGCGGAACAGCTCT  5640

Qy  5641  TGGTACATGGGTGTGGTTTGTCTTGTTGTCTTTGACATGCGTTAGGGCTGCCATGCGTCCGCCTC  5700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641  TGGTACATGGGTGTGGTTTGTCTTGTTGTCTTTGACATGCGTTAGGGCTGCCATGCGTCCGCCTC  5700

Qy  5701  ATCCACTGCTTGCTTGACATTTAAATGCTTGACATTTAAATGCTTGTGAGTGGGTGAGTGGTCGCCACTATGGATCAGCT  5760


Qy  5701  ATCCACTGCTTGCTTGACATTTAAATGCTTGACATTTAAATGGGTGAGTGAGTGGTCGCCACTATGGATCAGCT  5760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701  ATCCACTGCTTGCTTGACATTTAAATGCTTGACATTTAAATGGGTGAGTGAGTGGTCGCCACTATGGATCAGCT  5760

Qy  5761  TGCTGGTTAGTCTACTCGTCCGGTTCAATCCGGTTCAATCCGGCCGCGGAGTTGTGGGCGTCTTGTGCAGC  5820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761  TGCTGGTTAGTCTACTCGTCCGGTTCAATCCGGTTCAATCCGGCCGCGGAGTTGTGGGCGTCTTGTGCAGC  5820

Qy  5821  TTGTGCAATGTTGCTTGCTTGACAACAGCAGGGCCAGATCACTACTGGGCCCAACAGACTTCTTAC  5880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821  TTGTGCAATGTTGCTTGCTTGACAACAGCAGGGCCAGATCACTACTGGGCCCAACAGACTTCTTAC  5880

Qy  5881  TATGCTTGCTAGGAGCAACAACTCGTATGTAATGATGAGTACTTTATTGCCACTCGTGACATCCG  5940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881  TATGCTTGCTAGGAGCAACAACTCGTATGTAATGATGAGTACTTTATTGCCACTCGTGACATCCG  5940

Qy  5941  CAGGAAGATACTGGGCATTCGGAGGCATCTCACCCCTGGAGTGTCATATCAGCTTGCAT  6000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941  CAGGAAGATACTGGGCATTCGGAGGCATCTCACCCCTGGAGTGTCATATCAGCTTGCAT  6000

Qy  6001  CCGTTGGCTCCACACCCGACGGAGGATGAGTGATTGATTGCGGCCTCATTGCTTCAGAGAT  6060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001  CCGTTGGCTCCACACCCGACGGAGGATGAGTGATTGATTGCGGCCTCATTGCTTCAGAGAT  6060

Qy  6061  TTGGCAGTATGTGCAATTCTTTGTGATTTGCTTTGCTTTAATGTCCTTAAAGCTGAGTTCA  6120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061  TTGGCAGTATGTGCAATTCTTTGTGATTTGCTTTGCTTTAATGTCCTTAAAGCTGAGTTCA  6120

Qy  6121  GAGCATGTTAACATTCCTGGTTGTCCTTCTCACAGCTGCCAGAGCTGCCAGAGGGGCCCC  6180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121  GAGCATGTTAACATTCCTGGTTGTCCTTCTCACAGCTGCCAGAGCTGCCAGAGGGGCCCC  6180

Qy  6181  CTGATTGATCAGGTATGCTTCAAGCACGCTGTCCATGCGGTCTGAACTCATCTTTTC  6240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181  CTGATTGATCAGGTATGCTTCAAGCACGCTGTCCATGCGGTCTGAACTCATCTTTTC  6240

Qy  6241  TGTTGAGAATGGTTTGCAAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241  TGTTGAGAATGGTTTGCAAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300

Qy  6301  AGGGGCTGTTCAACGTGGCCTGGCTGTGTCGGTCGGCGGACCGGACCCAACTGATTG  6360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301  AGGGGCTGTTCAACGTGGCCTGGCTGTGTCGGTCGGCGGACCGGACCCAACTGATTG  6360

Qy  6361  GACTAGTCTTGTCGTCAATTATGCGTCAATTATGCGTAAATATGAGAAATGGGAGA  6420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361  GACTAGTCTTGTCGTCAATTATGCGTCAATTATGCGTAAATATGAGAAATGGGAGA  6420

Qy  6421  TCACATTTTGTTGTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCAAC  6480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421  TCACATTTTGTTGTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCAAC  6480

Qy  6481  CTTGAGAGCTGCACGTGGCCGTGGCCATGGCCGTGTTATCTAGTGTAATCTAGTGAGCCCAA  6540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481  CTTGAGAGCTGCACGTGGCCGTGGCCATGGCCGTGTTATCTAGTGTAATCTAGTGAGCCCAA  6540

Qy  6541  TGCACTTGAGACAAATGACTGTAATTCCACAAACAACACTCCTAGTGATGAAGCCGCAGT  6600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541  TGCACTTGAGACAAATGACTGTAATTCCACAAACAACACTCCTAGTGATGAAGCCGCAGT  6600

Qy  6601  GTCCGCTCTTGTTTCAAACAGGAGTTGCGCGTACAAACCAATTTGCTTGAGGCAATTTC  6660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601  GTCCGCTCTTGTTTCAAACAGGAGTTGCGCGTACAAACCAATTTGCTTGAGGCAATTTC  6660

Qy  6661  AGCTGGCGGTTGACACCAAACTGACAATTCCATCCGAGGAGTGAGTGGTCGCCCTCCCATCCCAAACTGCGTGAAGAAA  6720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661  AGCTGGCGGTTGACACCAAACTGACAATTCCATCCGAGGAGTGAGTGGTCGCCCTCCCATCCCAAACTGCGTGAAGAAA  6720

Qy  6721  AGCTGCGGTTGACATCGGTGGGTTTGTCTTTTGTTTTCAAACAGGACCACCAAACTGCGTGAAGAAA  6780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721  AGCTGCGGTTGACATCGGTGGGTTTGTCTTTTGTTTTCAAACAGGACCACCAAACTGCGTGAAGAAA  6780

Qy  6781  AGCTGCGGTTGACATCGGTGGGTTTGTCTTTTGTTTTCAAACAGGACCACCAAACTGCGTGAAGAAA  6840
```

```
Db   6781 AGCTGGCGTTGACACCACCAAACTGCCAGCCCCTCCATCGAAGAGGTAGTGGTAAGAAA 6840
                                                                         6900
Qy   6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTCGCCTCCCCTCCCCGGAGATCCGTCCC    6900
                |||||||||||||||||||||| |||||||||||| |||||||||||||||||||||
Db   6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTCGCCTGCCTCCCCTCCCCGGAGATCCGTCCC 6900

Qy   6901 AGGAGTGTCATGTCCTGAACGAAGTGACCCGTTAGAGGTCTTCAAACCT            6960
                ||||||||||||||||||||||||||||||||||||||||||||||||
Db   6901 AGGAGTGTCATGTCCTGAACGAAGTGACCCGTTAGAGGTCTTCAAACCT 6960

Qy   6961 CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGGCGATGCCCTGTTGGGAGCCGGGTGA 7020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6961 CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGGCGATGCCCTGTTGGGAGCGGGGTGA 7020

Qy   7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGCGAAACAGGCGGAGGCCCTGATGA   7080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7021 GTGTAACCGTTTCACTGCAATTGGATGTGCAATGCGAAACAGGCGGAGGCCCTGATGA 7080

Qy   7081 TTTACCCAGTTACCCTCCCAAAAAGAGAGGTCTCTGAATGTCAGACGAAAGTTGGTCGAC 7140
                ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   7081 TTTACCCAGTTACCCTCCCAAAAAGAGAGGTCTCTGAATGTCAGACGAAAGTTGGTCGAC 7140

Qy   7141 GGCTACAACCGCTTCAGCTACGTTACTGGCCCCCGTACCCTAGAGATACCGCGGAAAGGA 7200
                ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   7141 GGCTACAACCGCTTCAGCTACGTTACTGGCCCCCGTACCCTAGAGATACCGCGGAAAGGA 7200

Qy   7201 TTTCCACTCAGTCAGCCCCCGCCAAACGGCCTACAAAAAGAAGTTGGAAAGAGTGAGTT 7260
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7201 TTTCCACTCAGTCAGCCCCCGCCAAACGGCCTACAAAAAGAAGTTGGAAAGAGTGAGTT 7260

Qy   7261 TTCGTGCAGCATGAGCTGACCTGACCGACGGCAGATTAGCTTCAAAACTGCTTCTAAAGT 7320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7261 TTCGTGCAGCATGAGCTGACCTGACCGACGGCAGATTAGCTTCAAAACTGCTTCTAAAGT 7320

Qy   7321 TCTGTCTGCAACTCGGGCCATCACTAGTGTTCCTCAAACAAGATCATTGGTTCATTGGT  7380
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7321 TCTGTCTGCAACTCGGGCCATCACTAGTGTTCCTCAAACAAGATCATTGGTTCATTGGT 7380

Qy   7381 GACTGAGCCGCGGGAGATGCGGAGCTTAGAGAAACAAAAGTCACTATTAATAGACAACCTCT 7440
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7381 GACTGAGCCGCGGGAGATGCGGAGCTTAGAGAAACAAAAGTCACTATTAATAGACAACCTCT 7440

Qy   7441 GTTCCCCCATCATACCACACAAGCAGCAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7441 GTTCCCCCATCATACCACACAAGCAGCAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500

Qy   7501 CGGTGTCATGTGGGAGGAGTCTTCGTGAAGCAGCAGTAGCAGCTCACACGCGCCTCTAAGTC 7560
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7501 CGGTGTCATGTGGGAGGAGTCTTCGTGAAGCAGCAGTAGCAGCTCACACGCGCCTCTAAGTC 7560

Qy   7561 CCACATCACTGCCTTCGGGGCACTGATGTTCGTTCTCGGAGCAGCCCGCAAGGCTGTTCT 7620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7561 CCACATCACTGCCTTCGGGGCACTGATGTTCGTTCTCGGAGCAGCCCGCAAGGCTGTTCT 7620

Qy   7621 GGACTTGCAGAAGTGTGTCGGAGGAGTGCGAGGATATCGGCAGAGTTCATTATCCGTGAT 7680
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7621 GGACTTGCAGAAGTGTGTCGGAGGAGTGCGAGGATATCGGCAGAGTTCATTATCCGTGAT 7680

Qy   7681 AGTTCCAAAGGAGGAGTCTTCGTGAAGACTATGGAAATGAGAACCAACAACAAAGAAACCCCAAG 7740
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7681 AGTTCCAAAGGAGGAGTCTTCGTGAAGACTATGGAAATGAGAACCAACAACAAAGAAACCCCAAG 7740

Qy   7741 GCTTATCTCGTTACCCCCACCTTGTTCGTTCTCGGAGATGTGTTGAGAGATGTTACTACGGTCAGGT 7800
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7741 GCTTATCTCGTTACCCCCACCTTGTTCGTTCTCGGAGATGTGTTGAGAGATGTTACTACGGTCAGGT 7800

Qy   7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCCGGTTTGTAGATCCACGTAC       7860
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCCGGTTTGTAGATCCACGTAC 7860

Qy   7861 CCGTGTCAAGGCTCTGTTGTCGATGTGTCACCCGATGCAGTCGGAGCCACATGCCGATAC 7920
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7861 CCGTGTCAAGGCTCTGTTGTCGATGTGTCACCCGATGCAGTCGGAGCCACATGCCGATAC 7920
```

```
Qy   7921 AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGTGGGAGACAGACATCTACTC  7980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7921 AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGTGGGAGACAGACATCTACTC 7980

Qy   7981 AGCAGCTAAACTCAGTGACCAACACCGACCAACACCGACGCTGCATTCACACCATTGCGAGGCAGTTATA 8040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7981 AGCAGCTAAACTCAGTGACCAACACCGACCAACACCGACGCTGCATTCACACCATTGCGAGGCAGTTATA 8040

Qy   8041 CGCTGAGGACCGATGATCGGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTAGGTC 8100
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8041 CGCTGGAGGACCGACCGATGATCGGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTAGGTC 8100

Qy   8101 TTCCGGCGTCTATACTACCCAAGTTCCAACAGTTTGACCTTGACCTGCTGCGGCTGAAGTAAATGC 8160
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8101 TTCCGGCGTCTATACTACCCAAGTTCCAACAGTTTGACCTTGACCTGCTGCGGCTGAAGTAAATGC 8160

Qy   8161 TGCAGCCGAACAGGCTGGCATGAAGAACCCTCGCTTCCTTATTTGCCGCGATGATTGCAC 8220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8161 TGCAGCCGAACAGGCTGGCATGAAGAACCCTCGCTTCCTTATTTGCCGCGATGATTGCAC 8220

Qy   8221 CGTAATTTGGAAGAGCGGCCGAGCAGATGCAGACAAACAAGCAATGCCGTGTCTTTGCTAG 8280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8221 CGTAATTTGGAAGAGCGGCCGAGCAGATGCAGACAAACAAGCAATGCCGTGTCTTTGCTAG 8280

Qy   8281 CTGGATGAAGGTGATGGGTGCACCACAGATTGTGCCTCAACCCAAATACAGTTTGA    8340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8281 CTGGATGAAGGTGATGGGTGCACCACAGATTGTGCCTCAACCCAAATACAGTTTGA 8340

Qy   8341 AGAATTAACATCATGTCATCAAATGTTACCTCTGGAATTACGTGGCAAGCCTTA       8400
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8341 AGAATTAACATCATGTCATCAAATGTTACCTCTGGAATTACGTGGCAAGCCTTA 8400

Qy   8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCTGGCAGGTGCTCTGCCGAGGGTCTCGGG 8460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCTGGCAGGTGCTCTGCCGAGGGTCTCGGG 8460

Qy   8461 ATACAACCCCAGTGCTGCTGCGGTGATTGGGTATCGGATTCTAATACATCACTACCCATGTTTGTGGGT 8520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8461 ATACAACCCCAGTGCTGCTGCGGTGATTGGGTATCGGATTCTAATACATCACTACCCATGTTTGTGGGT 8520

Qy   8521 TAGCCGTGTGTGTGGCGTGTCCATTTCATTGGAGCAGAGATGCTCTTTGAGGACAAACTTCCCGA 8580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8521 TAGCCGTGTGTGTGGCGTGTCCATTTCATTGGAGCAGAGATGCTCTTTGAGGACAAACTTCCCGA 8580

Qy   8581 GACTGTGACCTTGACTGGTATGATGGAAAAATTATACGGTGCCTGTAGAGATCGTGCCCAG 8640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8581 GACTGTGACCTTGACTGGTATGATGGAAAAATTATACGGTGCCTGTAGAGATCGTGCCCAG 8640

Qy   8641 CATCATTGCTGGTGTGCACGGTATTGAGGCTTTCTCGGTGCGGTGCGCTACACCAACGCTGA 8700
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8641 CATCATTGCTGGTGTGCACGGTATTGAGGCTTTCTCGGTGCGGTGCGCTACACCAACGCTGA 8700

Qy   8701 GATCCTCAGAGTTTCCAATCACTAACGACTTCATTATCGGCAAACTTGTGAT        8760
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8701 GATCCTCAGAGTTTCCAATCACTAACGACTTCATTATCGGCAAACTTGTGAT 8760

Qy   8761 AAAGAAAGCCAGGCGCGTCCTCGCCAGCGCCAAGAGGCGTGGGCGGAGCCACACGCAAATT 8820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8761 AAAGAAAGCCAGGCGCGTCCTCGCCAGCGCCAAGAGGCGTGGGCGGAGCCACACGCAAATT 8820

Qy   8821 GGCTCGCTTCCTTCTCGCGATGCCGATCTACATCTAGACCTTCCCGGAGGGGATTGGATAAGACGAG 8880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8821 GGCTCGCTTCCTTCTCGCGATGCCGATCTACATCTAGACCTTCCCGGAGGGGATTGGATAAGACGAG 8880

Qy   8881 CGTGGCTCGGTACCACCTTCAATTATTGTGATGTTTACTCCCCGGAGGGGATGTGTT    8940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8881 CGTGGCTCGGTACCACCTTCAATTATTGTGATGTTTACTCCCCGGAGGGGATGTGTT 8940

Qy   8941 TATTACCCACAGAGAAGAGATTGCCAGAAGTTCCTTGTGAAGTATTTGGCTGGTGTCATTGTTTT 9000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8941 TATTACCCACCACAGAGAAGAGATTGCCAGAAGTTCCTTGTGAAGTATTTGGCTGGTGTCATTGTTTT 9000
```

Qy   9001 TGCCCTAGGGCTCATTGCTGTGTTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA 9060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9001 TGCCCTAGGGCTCATTGCTGTGTTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA 9060

Qy   9061 CTAACAGTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGCAACAGGGGAGACCCCCGGGC 9120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9061 CTAACAGTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGCAACAGGGGAGACCCCCGGGC 9120

Qy   9121 TTAACGACCCGCGATGTGAGTTGGGGACCATGGTGGATCAGAACCGTTTCGGGGTGAA 9180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9121 TTAACGACCCCGCGATGTGAGTTGGCGACCATGGTGGATCAGAACCGTTTCGGGGTGAA 9180

Qy   9181 GCCATGTGCTCGAAGGGGATGACGCTCCCTTCTGCTCATCCACAAAAACCGTCTCGGGTGG 9240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9181 GCCATGTGCTCGAAGGGGATGACGCTCCCTTCTGCTCATCCACAAAAACCGTCTCGGGTGG 9240

Qy   9241 GTGAGGAGTCCTGGCTGTGTGGGAAGCAGTCAGTATAATTCCCGTCGTGTGTGGTGACGC 9300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9241 GTGAGGAGTCCTGGCTGTGTGGGAAGCAGTCAGTATAATTCCCGTCGTGTGTGGTGACGC 9300

Qy   9301 CTCACGACGTATTTGTCCGCTGCGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTG 9360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9301 CTCACGACGTATTTGTCCGCTGCGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTG 9360

Qy   9361 TTCCAAGCGGAGGGCAACCCCCGCTTGGAATTAAAAACT 9399
           |||||||||||||||||||||||||||||||||||||||
Db   9361 TTCCAAGCGGAGGGCAACCCCCGCTTGGAATTAAAAACT 9399

RESULT 2
AF179612
LOCUS   AF179612   9399 bp   RNA   linear   VRL 07-OCT-1999
DEFINITION Hepatitis GB virus B polyprotein gene, complete cds.
ACCESSION   AF179612
VERSION   AF179612.1 GI:6014504
KEYWORDS   .
SOURCE   Hepatitis GB virus B
  ORGANISM Hepatitis GB virus B
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae.
REFERENCE   1 (bases 1 to 9399)
  AUTHORS   Bukh,J., Apgar,C.L. and Yanagi,M.
  TITLE   Toward a surrogate model for hepatitis C virus: An infectious
            molecular clone of the GB virus-B hepatitis agent
  JOURNAL   Virology 262 (2), 470-478 (1999)
   MEDLINE   99434247
   PUBMED   10502525
REFERENCE   2 (bases 1 to 9399)
  AUTHORS   Bukh,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (23-AUG-1999) Hepatitis Viruses Section, Laboratory of
            Infectious Diseases, National Institute of Allergy and Infectious
            Diseases, National Institutes of Health, Building 7, Room 201, 7
            Center Dr. 0740, Bethesda, MD 20892, USA
FEATURES   Location/Qualifiers
     source   1..9399
                     /organism="Hepatitis GB virus B"
                     /mol_type="genomic RNA"
                     /specific_host="unknown"
                     /db_xref="taxon:39113"
                     /note="most closely related to hepatitis C virus;
                     experimental infection of tamarins causes acute resolving
                     hepatitis"
     CDS   446..9040
                     /codon_start=1
                     /product="polyprotein"
                     /protein_id="AAF01368.1"
                     /db_xref="GI:6014505"
                     /translation="MPVISTGTTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKV
                     QRDARPRNYKIAGIHDGLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVT
                     THTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVT
                     DPDTNVTILTNCCQRNQVIVCSPSTCLHEPGVICADECWVPANPYVISHPSNWTGTDS
                     FLADHIDWFVMGALVTCDALDIGELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTG
                     IDPGPLGFIGWMAGKVEAVIPITKLASQVPVAIATMFSSVHVLAVGALIYYASRGKWY

ORIGIN

Query Match   100.0%; Score 9399; DB 14; Length 9399;
Best Local Similarity   100.0%;
Matches 9399; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACCACAAACACTCCAGTTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCACCTAGTAT   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCACAAACACTCCAGTTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCACCTAGTAT   60

Qy   61 CAGGGCGTGGGGGATTTCCCCTGCCGTCTCGAGAAGGTGGAGCCAACCACCACTTAGTAT   120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CAGGGCGTGGGGGATTTCCCCTGCCGTCTCGCAGAAGGTGGAGCCAACCACCACTTAGTAT   120

Qy   121 GTAGGCGGCGGCGGCGGCGGCGTTCGGTGGTCGGTGGTGGGCGGCTTTAGGCAGCCTCCACGGCCACCA   180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTAGGCGGCGGCGGCGGCGGCGTTCGGTGGTCGGTGGTGGGCGGCTTTAGGCAGCCTCCACGGCCACCA   180

Qy   181 CCTGATGGGCGGTTCATGGGTTCGGTGGTGGTGGGCGGCTTTAGGCAGCCTCCACGCCCACCA   240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCTGATGGGCGGTTCATGGGTTCGGTGGTGGTGGGCGGCTTTAGGCAGCCTCCACGCCCACCA   240

Qy   241 CCTCCCAGATAGAGCGGCGGCCACTGTAGGGAAGACCGGGGACCGGTCACTACCAAGGACG   300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCCCAGATAGAGCGGCGGCCACTGTAGGGAAGACCGGGGACCGGTCACTACCAAGGACG   300

Qy   301 CAGACCTCTTTTGAGTATCACGCTCCGGAAGTAGTTGGGCAAGCCACCTATATGTGT   360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGACCTCTTTTGAGTATCACGCTCCGGAAGTAGTTGGGCAAGCCACCTATATGTGT   360

```
Qy  1441 GTGGTATCAGTTGCTCCTAGGCTTATGCTTTACATAGAAGCGACCTCTGGAAACCCCAT 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 GTGGTATCAGTTGCTCCTAGGCTTATGCTTTACATAGAAGCGACCTCTGGAAACCCCAT 1500

Qy  1501 CAGGGTGCCCACTGATGCTCAATAGCTGAGTTTGCTCGCGTTGCTCGCCTTTGCTCGGC 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 CAGGGTGCCCACTGATGCTCAATAGCTGAGTTTGCTCGCGTTTGCTCGCCATGTCC 1560

Qy  1561 TTGCCACTCTTATTTGAGTGAGAATGTCAGAAGTCATTTGTTACAGTCCAAAGTGGAC 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 TTGCCACTCTTATTTGAGTGAGAATGTCAGAAGTCATTTGTTACAGTCCAAAGTGGAC 1620

Qy  1621 CAGGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCTATACAATCCCTGG 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 CAGGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCTATACAATCCCTGG 1680

Qy  1681 TGCGAGGGGATGTATGGTAAATTCAAAATTAACACATGGGGTTGCTGCCGTATTCGCAA 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 TGCGAGGGGATGTATGGTAAATTCAAAATTAACACATGGGGTTGCTGCCGTATTCGCAA 1740

Qy  1741 TGTGCCATCGTACTGCACTATGGCACTGATGCAGTGTGGAACGACACTCGCAACACTTA 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 TGTGCCATCGTACTGCACTATGGCACTGATGCAGTGTGGAACGACACTCGCAACACTTA 1800

Qy  1801 CGAAGCATGCGGTAACAACCATGGCTAACACCGTAACAACCGCAGCAACGGCTCAGCCCTGAA 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 CGAAGCATGCGGTAACAACCATGGCTAACACCGTAACAACCGCAGCAACGGCTCAGCCCTGAA 1860

Qy  1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC 1920

Qy  1921 AGGCCATTGTATTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCTGTGAA 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 AGGCCATTGTATTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCTGTGAA 1980

Qy  1981 TTCCACTCTCCTACCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCACCTGTGGTACG 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 TTCCACTCTCCTACCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCACCTGTGGTACG 2040

Qy  2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100

Qy  2101 ATTGATCACCAAAGACAAAGACTGGAAAAATATCAGTCTTATATTCCGCCACGGGTGC 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 ATTGATCACCAAAGACAAAGACTGGAAAAATATCAGTCTTATATTCCGCCACGGGTGC 2160

Qy  2161 TTTGTCTCTTACGGAGTTACCACCAAGGCGTGGTGCTAATTCTCGTTGTGTGG 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 TTTGTCTCTTACGGAGTTACCACCAAGGCCGTGGTGCTAATTCTCGTTGTGTGG 2220

Qy  2221 CAGCAAGTATCTTATTTTAGCCTTTTAGCCTCGTTGTTCCTTGTTTCTTCAAGCTGGATTC 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 CAGCAAGTATCTTATTTTAGCCTTTTAGCCTCGTTGTTCCTTGTTTCTTCAAGCTGGATTC 2280

Qy  2281 TGGTTACCCTTGCGTCCTGTCTCCATCTTTGCTTTGATTTTCTTCATCTGTTCTCCG 2340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 TGGTTACCCTTGCGTCCTGTCTCCATCCCAGTCGTATTTTGATTTTCTTCATCTGTTCTCCG 2340

Qy  2341 TTTGTCTAAAGCTCAAGTAGTCCTGTATTGTTGCCCCTTTTTAGGGTTGTGCCCCCT 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 TTTGTCTAAAGCTCAAGTAGTCCTGTATGTTGCCCCTTTTTTAGGGTTGTGCCCCCT 2400

Qy  2401 CTGCAGGCTACGTTATGCGCCATATGGCCCACAATGGCCTATTGCGACTAT 2460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 CTGCAGGCTACGTTATGCGCCATATGGCCCACAATGGCCTATTGCGACTAT 2460

Qy  2461 AACTTTCTTTGTTGTTGCAGCGTCTGCCACCAGATTATGACTGGTGGGTCGCGGGCAA 2520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 AACTTTCTTTGTTGTTGCAGCGTCTGCCACCAGATTATGACTGGTGGGTCGCGGGCAA 2520

Qy  2521 AGTGGCAGGGTTAGTTTTTGTTGGGCCGGCCGTAACCGTGGTCACCGCATAGCTCTGCTTGT 2580
```

```
Qy   361 TGGGATGGTTGGGGGTTAGCCATCATCACTACTGCCTGATAGGGTCCTTGCGAGGGGAT 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TGGGATGGTTGGGGGTTAGCCATCATCACTACTGCCTGATAGGGTCCTTGCGAGGGGAT 420

Qy   421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACC 480

Qy   481 TGCGCCCAGAACGCGCCAAGAACAAGCAGCGCAGGCTTCATATCCTGTGTCCATTAAAAC 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TGCGCCCAGAACGCGCCAAGAACAAGCAGCGCAGGCTTCATATCCTGTGTCCATTAAAAC 540

Qy   541 ATCTGTTGAAAGGGGACAACAGCAGCAAAGCCAGCGCGGATGCGGCTCGGCTTCGGTAA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 ATCTGTTGAAAGGGGACAACAGCAGCAAAGCCAGCGCGGATGCGGCTCGGCTTCGGTAA 600

Qy   601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGCTCAGGCTCGCGCTGCCAGC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGCTCAGGCTCGCGCTGCCAGC 660

Qy   661 TCATGGTTGGGGACGCCAAGACCCTCGCGCAATAGTCTGGAATCCTTCTGGA 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TCATGGTTGGGGACGCCAAGACCCTCGCGCAATAGTCTGGAATCCTTCTGGA 720

Qy   721 TTACCCTTTGGGGGTTGGATTGGTGATGTTACAACTCACACACCCTTCAGTAGGCCCGCTGGT 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TTACCCTTTGGGGGTTGGATTGGTGATGTTACAACTCACACACCCTTCAGTAGGCCCGCTGGT 780

Qy   781 GGCAGAGCCGGTCGTTGCAGTACGGCGTTGCTGGGATAGCGGATGGAGTCAA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GGCAGAGCCGGTCGTTGCAGTACGGCGTTGCTGGGATAGCGGATGGAGTCAA 840

Qy   841 CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTGTGGTATGTCTGCTATCTTTCGGCCTG 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTGTGGTATGTCTGCTATCTTTCGGCCTG 900

Qy   901 TCCCTGTAGTGGGGCCGCGGGTCACTGACCAGACACAAATACCACCAAATTG 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCCTGTAGTGGGGCCGCGGGTCACTGACCAGACACAAATACCACCAAATTG 960

Qy   961 CTGCCAGCGTAATCAGGTTATCTATTGTTCCTTGCCTACACGAGCCTGGTTG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CTGCCAGCGTAATCAGGTTATCTATTGTTCCTTGCCTACACGAGCCTGGTTG 1020

Qy  1021 TGTGATCTGTGCGGACGAGTGCTGGTTCCCGCCAATCCGTACACTCACACCCTTCCAA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 TGTGATCTGTGCGGACGAGTGCTGGTTCCCGCCAATCCGTACACTCACACCCTTCCAA 1080

Qy  1081 TTGGACTGGCACGACTCCTTCTTGGCTGACCACATTGATTTGTTAGTGTGCTCTTGT 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TTGGACTGGCACGACTCCTTCTTGGCTGACCACATTGATTTGTTAGTGTGACTGT 1140

Qy  1141 GACCTGTGCAGCGCCCCTTGACATTGGTGTGTGCGTGTTGTTAGTGCGGTTGACTG 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 GACCTGTGCAGCGCCCCTTGACATTGGTGTGTGCGTGTTGTTAGTTGCGGTTGACTG 1200

Qy  1201 GCTTGTCAGGCACTGGAATAGATCCTGGTTATTCACATGAAACTGGTACTTGTTACCTGGA 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GCTTGTCAGGCACTGGAATAGATCCTGGTTATTCACATGAAACTGGTACTTGTTACCTGGA 1260

Qy  1261 AGTGGGCACTGGAATGGTTTATCGGGTGGTTATCGGGTGGATGGCCGGCCAAGGT 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 AGTGGGCACTGGAATGGTTTATCGGGTGGTTATCGGGTGGATGGCCGGCCAAGGT 1320

Qy  1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT 1380

Qy  1381 GTTTAGCAGTGTACACTACCTGGCCGGTTGGCGCTCTGATCTACTACTATGCCTCTCGGGGCAA 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 GTTTAGCAGTGTACACTACCTGGCCGGTTGGCGCTCTGATCTACTACTATGCCTCTCGGGGCAA 1440
```

```
Db  2521  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  2580
          AGTTGCAGGGTTAGTTTTGTGGGCGCCGGCCGGCCGTAACCGTGGTCACCGCATAGCTCTGCTTGT

Qy  2581  AGGTCCTTGGCCTCGGTAGGCGCTTTAACCCTCGTGCATTTGGTTACGCCTGCTTCAGC  2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  AGGTCCTTGGCCTCGGTAGGCGCTTTAACCCTCGTGCATTTGGTTACGCCTGCTTCAGC  2640

Qy  2641  TTTGATACCGGAGATAATTGGAGAGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT  2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  TTTGATACCGGAGATAATTGGAGAGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT  2700

Qy  2701  GTCTCGTTTGGCTTCTTGCTCACTGTTACCTCGCGTGCTGTTAGTTAGTTAACTCCTATCT  2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  GTCTCGTTTGGCTTCTTGCTCACTGTTACCTCGCGTGCTGTTAGTTAGTTAACTCCTATCT  2760

Qy  2761  TTGGCAACGTTGGGAGAATTGGTTTGGAACGTTACACTAAGACCGGAGAGGTTTTTCCT  2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  TTGGCAACGTTGGGAGAATTGGTTTGGAACGTTACACTAAGACCGGAGAGGTTTTTCCT  2820

Qy  2821  TGTGCTGGTTGTTCCCCGGTGCGACATATGACGCGCTGGTGACTTTCTGTCGTGTGTCA  2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  TGTGCTGGTTGTTCCCCGGTGCGACATATGACGCGCTGGTGACTTTCTGTCGTGTGTCA  2880

Qy  2881  CGTAGCTCTCTATGTTTAACATCCAGTGCAGCAGTCGTTCTTTGGGACTGACTCTAGGGT  2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  CGTAGCTCTCTATGTTTAACATCCAGTGCAGCAGTCGTTCTTTGGGACTGACTCTAGGGT  2940

Qy  2941  TAGGGCCCATAGAAGTGGTGCGCTCTCGGAAGTGTGCGCTTCGGTATTCTCATTATGT  3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941  TAGGGCCCATAGAAGTGGTGCGCTCTCGGAAGTGTGCGCTTCGGTATTCTCATTATGT  3000

Qy  3001  TCTTAAGTTTTCCTCTTAGTGTTGGTGTAGAATGGTGTGTTTTTCTATAAGGACTTGCA  3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  TCTTAAGTTTTCCTCTTAGTGTTGGTGTAGAATGGTGTGTTTTTCTATAAGGACTTGCA  3060

Qy  3061  TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCGCAAGAGCCATTTTTCCC  3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061  TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCGCAAGAGCCATTTTTCCC  3120

Qy  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGAAGACGCTTGGGCGTGTGGGGACAC  3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGAAGACGCTTGGGCGTGTGGGGACAC  3180

Qy  3181  GGTTGATGGTTTGTTGCCCGTTGTGCGCGTCTCGGCGACCTTGTTTTCGGCGTCTTGGCTAT  3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181  GGTTGATGGTTTGTTGCCCGTTGTGCGCGTCTCGGCGACCTTGTTTTCGGCGTCTTGGCTAT  3240

Qy  3241  GCCGCCAGATGGGTGGGCCATTACCGCACCTTTACGCTGCAGTGTCTCTCTGAACGTGG  3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241  GCCGCCAGATGGGTGGGCCATTACCGCACCTTTACGCTGCAGTGTCTCTCTGAACGTGG  3300

Qy  3301  CACGCTGTCAGCGGATGCAGCAGTGTCATGACTGGTATAGACCCCGAACTTGGACTGAAC  3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301  CACGCTGTCAGCGGATGCAGCAGTGTCATGACTGGTATAGACCCCGAACTTGGACTGAAC  3360

Qy  3361  TATCTTCAGATTAGGATCTCTGGCCACTCTCGGCCACTACATGGGATTGTTGTTGTGACAACGTGTT  3420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361  TATCTTCAGATTAGGATCTCTGGCCACTCTCGGCCACTACATGGGATTGTTGTTGTGACAACGTGTT  3420

Qy  3421  GTATACTGCTCACCATGCAGCAGAGGGGCGCAAGGGCGCCGGCCGGTTGCCTCATCCCAACAGGCTCTATACA  3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421  GTATACTGCTCACCATGCAGCAGAGGGGCGCAAGGGCGCCGGCCGGTTGCCTCATCCCAACAGGCTCTATACA  3480

Qy  3481  CCCAATAACCGTTGACGCGGCTAATGCACCAGGACATCTATCAACCACCATGTGGAGCTGG  3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481  CCCAATAACCGTTGACGCGGCTAATGCACCAGGACATCTATCAACCACCATGTGGAGCTGG  3540

Qy  3541  GTCCCTTACTCGGTGCTCTTGCGGGAGACCAAGGGTATCTGGTAACGACTGGGGTC  3600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541  GTCCCTTACTCGGTGCTCTTGCGGGAGACCAAGGGTATCTGGTAACGACTGGGGTC  3600

Qy  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGCGGGGCCCTTCCCAT  3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGCGGGGCCCTTCCCAT  3660

Qy  3661  GGCTGTTGCCAAGGGTTCTTCAGGTGCCCCGATTCGTGCTCCTCCGGGCATGTTATTGG  3720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661  GGCTGTTGCCAAGGGTTCTTCAGGTGCCCCGATTCGTGCTCCTCCGGGCATGTTATTGG  3720

Qy  3721  GATGTTCACCGCTGCTAGAAATTCTCGGCGGTTCAGTCAGTTAGGGTTAGGCCGGTT  3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721  GATGTTCACCGCTGCTAGAAATTCTCGGCGGTTCAGTCAGTTAGGGTTAGGCCGGTT  3780

Qy  3781  GGTGTGTGCTGGATACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC  3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781  GGTGTGTGCTGGATACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC  3840

Qy  3841  TGTGCCTAACGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC  3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841  TGTGCCTAACGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC  3900

Qy  3901  CAAATTACCACTTTCTTACATGCAGGAGAAGTATGAGGTCTTCGGTTCCTAAATCCCAGTGT  3960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901  CAAATTACCACTTTCTTACATGCAGGAGAAGTATGAGGTCTTCGGTTCCTAAATCCCAGTGT  3960

Qy  3961  GGCTACAACAGCATCAATGCCAAAGTACATGCCGACGTACGGCGTAGGGCGTGAATCCAAATTG  4020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961  GGCTACAACAGCATCAATGCCAAAGTACATGCCGACGTACGGCGTAGGGCGTGAATCCAAATTG  4020

Qy  4021  CTATTTTAATGGCAAATGTACCAACACAGGGGCTCACTACGTACGACACATATGGCAT  4080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021  CTATTTTAATGGCAAATGTACCAACACAGGGGCTCACTACGTACGACACATATGGCAT  4080

Qy  4081  GTACCTGACCGGACGCATGTTCCCGGAACTATGATGTAATCATTGTGACGAATGCCATGC  4140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081  GTACCTGACCGGACGCATGTTCCCGGAACTATGATGTAATCATTGTGACGAATGCCATGC  4140

Qy  4141  TACCGATGCAACCACCGTGTGGGCATTGGAAAGGTCCATCGAAAGTCCATCCAAAAA  4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141  TACCGATGCAACCACCGTGTGGGCATTGGAAAGGTCCATCGAAAGTCCATCCAAAAA  4200

Qy  4201  TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCTGAGTAATCCCCTTCATGGAAAAAAGAT  4260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201  TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCTGAGTAATCCCCTTCATGGAAAAAAGAT  4260

Qy  4261  CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGGAGAAAAGACACTG  4320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261  CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGGAGAAAAGACACTG  4320

Qy  4321  TAAGGAGAAAATCTGAAGAAAGGGAGACACCTTATCTTTGAGGCTACCAAAAACACTG  4380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321  TAAGGAGAAAATCTGAAGAAAGGGAGACACCTTATCTTTGAGGCTACCAAAAACACTG  4380

Qy  4381  TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG  4440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381  TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG  4440

Qy  4441  ATGTGACATCTCAAAAATCCCTGAGGGGGCACTGTGTAGTAGTTGCCACTGATGCCTTGTG  4500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441  ATGTGACATCTCAAAAATCCCTGAGGGGGCACTGTGTAGTAGTTGCCACTGATGCCTTGTG  4500

Qy  4501  TACAGGGTACACTGGTGACTTTGATTGTGTATGACTGGCAGCCTCATGGTGAAGGCAC  4560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501  TACAGGGTACACTGGTGACTTTGATTGTGTATGACTGGCAGCCTCATGGTGAAGGCAC  4560

Qy  4561  ATGCCATGTTGACCTTGACCTACTTTCACCATGGGAGCTGG  4620
          |||||||||||||||||||||||||||||||||||||||||
Db  4561  ATGCCATGTTGACCTTGACCTACTTTCACCATGGGAGCTGG  4620

Qy  4621  AATAGTTAAAGGCCAGCCGGCCGCACAGGCCGTGGGAGAGCTGGCATATATACTACTA  4680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621  AATAGTTAAAGGCCAGCCGGCCGCACAGGCCGTGGGAGAGCTGGCATATATACTACTA  4680

Qy  4681  TGTAGACCGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT  4740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681  TGTAGACCGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT  4740
```

Qy   5821 TTGTGCAATGTTGCTTTGACAACAGCAGAGGGCCAGATCACTGCCCAACAGACTTCTTAC 5880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5821 TTGTGCAATGTTGCTTTGACAACAGCAGAGGGCCAGATCACTGCCCAACAGACTTCTTAC 5880

Qy   5881 TATGCTTGCTAGGAGCAACAACTGTATGTGATGTAGTACTTATTGCCACTCGTGACATCCG 5940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5881 TATGCTTGCTAGGAGCAACAACTGTATGTGATGTAGTACTTTATTGCCACTCGTGACATCCG 5940

Qy   5941 CAGGAAGATACTGGCATTCTCGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT 6000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5941 CAGGAAGATACTGGCATTCTCGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT 6000

Qy   6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGCCTCATTGCTTGGGGTCTAGAGAT 6060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGCCTCATTGCTTGGGGTCTAGAGAT 6060

Qy   6061 TTGGCAGTATGTGTGCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGCTGGAGTTCA 6120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6061 TTGGCAGTATGTGTGCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGCTGGAGTTCA 6120

Qy   6121 GAGACATGGTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC 6180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6121 GAGCATGGTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC 6180

Qy   6181 CTGGATGGATCAGGTATGCTCCAAGCACGCTGTCCATGCGGTGCTGAACTCATCTTTTC 6240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6181 CTGGATTGGATCAGGTATGCTCCAAGCACGCTGTCCATGCGGTGCTGAACTCATCTTTTC 6240

Qy   6241 TGTTGAGAAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG 6300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6241 TGTTGAGAAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG 6300

Qy   6301 AGGGGCGTGTTCCAGTCACGCTAGGCTGTGTGTGGCTAGGCACCGATTG 6360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6301 AGGGGCGTGTTCCAGTCACGCTAGGCTGTGTGGCTAGGCACCGATTG 6360

Qy   6361 GACTAGTCCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA 6420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6361 GACTAGTCCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA 6420

Qy   6421 TCACATTTTGTTACAGCAGTATCCCTCTCCAAATGTCTGTTTCACCCGGTGCGCCCCCAAC 6480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6421 TCACATTTTGTTACAGCAGTATCCCTCTCCAAATGTCTGTTTCACCCGGTGCGCCCCCAAC 6480

Qy   6481 CTTGGAGAGCTGCAGTGGCCGTGACGGCGTTCAGTGTTATCTAGGTGAGCCCAA 6540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6481 CTTGGAGAGCTGCAGTGGCCGTGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540

Qy   6541 AACTCCTTGGACGACATCTGCTTGCTGACGGTCCTGACGGTAAGGGTAAAACTGTTAA 6600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6541 AACTCCTTGGACGACATCTGCTTGCTGACGGTCCTGACGGTAAGGGTAAAACTGTTAA 6600

Qy   6601 GCTTCCCTTCCCGCGTTGACGCGTCACACACCTGGTGCCGCATGCCAACTTAATTTGCGTGA 6660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6601 GCTTCCCTTCCCGCGTTGACGCGTCACACACCTGGTGCCGCATGCCAACTTAATTTGCGTGA 6660

Qy   6661 TGGACTTGAGACAAATGACTGTAATTCCACAAACAACACTCCTAGTGATGAAGCCGCAGT 6720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6661 TGGACTTGAGACAAATGACTGTAATTCCACAAACAACACTCCTAGTGATGAAGCCGCAGT 6720

Qy   6721 GTCCGCTCTTGTTTGTTTTCAAAACAGGAGTGCGCGGGTACAAAACCAATTGCTTGAGGCAATTTC 6780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6721 GTCCGCTCTTGTTTGTTTTCAAAACAGGAGTGCGCGGGTACAAAACCAATTGCTTGAGGCAATTTC 6780

Qy   6781 AGCTGGCGTTGACACCACCAAACTGCCAGCCCCCTCCATCGGAGAGTAGTGGTAAGAAA 6840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6781 AGCTGGCGTTGACACCACCAAACTGCCAGCCCCCTCCATCGGAGAGTAGTGGTAAGAAA 6840

Qy   6841 GCGCCAGTTCCGGCCAAGAACTGGTTCGTTCGCTTGCCTTCCTCCCCCCTCCGAGATCCGTCCC 6900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6841 GCGCCAGTTCCGGCCAAGAACTGGTTCGTTCGCTTGCCTTCCTCCCCCTCCGAGATCCGTCCC 6900

Qy   6901 AGGAGTGTCATGTCCTGCAACGAAGTGCCTGCAACGAAGTGCTGAAAGCCTTAGAGAAGTCCTTCAAAACCT 6960

Qy   4741 CGACGCAGCCAAGGCCATGGTATGGTTGTGTTCATCAACAGAAGCTCAAAACTATTCTGACAC 4800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 CGACGCAGCCAAGGCCATGGTATGGTTTGTGTTCATCAACAGAAGCTCAAAACTATTCTGACAC 4800

Qy   4801 CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTGGACGAGTGGGCTGATCT 4860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT 4860

Qy   4861 CTTTTCTATGGTCAACCCGAACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA 4920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861 CTTTTCTATGGTCAACCCGAACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA 4920

Qy   4921 TTATGTTTGTTGACTGCAGCCCAACTACAACTGTGTCATCAGTAGTATGGCTATGCTGCTCC 4980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 TTATGTTTGTTGACTGCAGCCCAACTACAACTGTGTCATCAGTAGTATGGCTATGCTGCTCC 4980

Qy   4981 CAATGACGCACCACGGTGGCAGGGAGCCCGGCTTGGGAAAAACCTTGTGGGGTTCTGTG 5040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 CAATGACGCACCACGGTGGCAGGGAGCCCGGCTTGGGAAAAACCTTGTGGGGTTCTGTG 5040

Qy   5041 GCGCATGGCTGACGGCGCTGACCGCCTGTCCTCGGCCCAGAGCCCAGCGGAGGTGACCAGATACCA 5100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041 GCGCTTGACGGCGCTGACCGCCTGTCCTCGGCCCAGAGCCCAGCGGAGGTGACCAGATACCA 5100

Qy   5101 AATGTGCTTCACTGAGAGTCAATACTTCTCGGGAACAGCGCGCCACTCGCTGTGTGGCGTGGAGT 5160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101 AATGTGCTTCACTGAGAGTCAATACTTCTCGGGAACAGCGCGCCACTCGCTGTGTGGCGTGGAGT 5160

Qy   5161 GGCTATGGCTTATCTAGCACTTTGACAACTTTGGCGGCCACTTGTGTGCGGCCGTTGCGTGGTC 5220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161 GGCTATGGCTTATCTAGCACTTTGACAACTTTGGCGGCCACTTGTGTGCGGCCGTTGCGTGGTC 5220

Qy   5221 TATTACATCAGTCCCTACTGTCGCGCCCCAGTGGTTGACGAAGAAGAAATCGT 5280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221 TATTACATCAGTCCCTACTGTCGCGCCCCAGTGGTTGACGAAGAAGAAATCGT 5280

Qy   5281 GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGGTTGCTTCATTGCGCAATTGACAAGCTGAA 5340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5281 GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGGTTGCTTCATTGCGCAATTGACAAGCTGAA 5340

Qy   5341 GAGTACAATCACCACAACTAGTTCCTTGAAAACCGCCTTGAAAAACTTAACAC 5400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5341 GAGTACAATCACCACAACTAGTTCCTTGAAAACCGCCTTGAAAAAACTTAACAC 5400

Qy   5401 CTTTCTTGGGCCTCATGCAGCTGCTACAATCCTTGCCTATCATGAGTATTGCTGTGTGGTTTAGT 5460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5401 CTTTCTTGGGCCTCATGCAGCTGCTACAATCCTTGCCTATCATGAGTATTGCTGTGTGGTTTAGT 5460

Qy   5461 CACTTTACCTGACAATCCCTTTGCATCATGCGTGTTTGCTTTCATTGCGGGTATTACTAC 5520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5461 CACTTTACCTGACAATCCCTTTGCATCATGCGTGTTTGCTTTCATTGCGGGTATTACTAC 5520

Qy   5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTTGGAGGCGCAATTGCGTCCAA 5580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTTGGAGGCGCAATTGCGTCCAA 5580

Qy   5581 GCTTACAGACGCTAGAGGGCCACTGGCGTTCATGATGGCCGGGGCTGCGGGAACAGCTCT 5640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5581 GCTTACAGACGCTAGAGGGCCACTGGCGTTCATGATGGCCGGGGCTGCGGGAACAGCTCT 5640

Qy   5641 TGGGTACGACATCGGTGGGTTTTGCTTCTTGACATGCTAGGCGGTATGCTGCGCCTC 5700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5641 TGGGTACGACATCGGTGGGTTTTGCTTCTTGACATGCTAGGCGGTATGCTGCGCCTC 5700

Qy   5701 ATCCACTGCTTGCTTGACATTTAAATGTTGATGGTGAGTGCCCACTATGGATCAGCT 5760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5701 ATCCACTGCTTGCTTGACATTTAAATGTTGATGGTGAGTGCCCACTATGGATCAGCT 5760

Qy   5761 TGCTGGTTTAGTCTACTCCCGGTTCAATCCGGCCGGGAGTTGTGGGCGTCTTGTTGTCAGC 5820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5761 TGCTGGTTTAGTCTACTCCCGGTTCAATCCGGCCGGGAGTTGTGGGCGTCTTGTTGTCAGC 5820

Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 8040
    AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCATTCACACCATTGCCGAGGCAGTTATA

Qy 7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCATTCACACCATTGCCGAGGCAGTTATA 8040

Db 8041 CGCTGGAGGACCGATGAGTCGCTTATGATGGCCGAGAGATCGGATATCGGATATGTGAGGTC 8100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 8041 CGCTGGAGGACCGATGAGTCGCTTATGATGGCCGAGAGATCGGATATCGGATATGTGAGGTC 8100

Db 8101 TTCCGGCGTCTATACTACCTCAAGTCCAACAGTTGACCTGCTTGGCTGAAGGTAAATGC 8160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 8101 TTTCGGCGTCTATACTACCTCAAGTCCAACAGTTGACCTGCTTGGCTGAAGGTAAATGC 8160

Db 8161 TGCAGCCGAACAGGCTGGCATGAAGAACCTCGCTTCCTTATTTGCGGCGATGATTGCAC 8220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 8161 TGCAGCCGAACAGGCTGGCATGAAGAACCTCGCTTCCTTATTTGCGGCGATGATTGCAC 8220

Db 8221 CGTAATTTGGAAGAGCGCCGGAGCAGATGCAGACAACAAGCAATGCGTGTCTTTGCTAG 8280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 8221 CGTAATTTGGAAGAGCGCCGGAGCAGATGCAGACAAACAAGCAATGCGTGTCTTTGCTAG 8280

Db 8281 CTGGATGAAGGTGATGGGTGCACCACAAGAGTGTGCTCAAATACAGTTTGGA 8340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 8281 CTGGATGAAGGTGATGGGTGCACCACAAGAGTGTGCTCAAATACAGTTTGGA 8340

Db 8341 AGAATTAACATCATGCTCATCAAATGTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 8341 AGAATTAACATCATGCTCATCAAATGTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400

Db 8401 CTACTTTCTTACAAGAGATCCTGTATCCCCCTGGCAGGTGCTCGCCGAGGGTCTCGGG 8460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 8401 CTACTTTCTTACAAGAGATCCTGTATCCCCCTGGCAGGTGCTCGCCGAGGGTCTCGGG 8460

Db 8461 ATACAACCCCAGTGCTGCGCTGGGTATTGGGTATCTAATAATACATCACTCTAATAGACAAACCTCT 8520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 8461 ATACAACCCCAGTGCTGCGCTGGGTATTGGGTATCTAATAATACATCACTCACTAATAGACAAACCTCT 8520

Db 8521 TAGCCGTGTGTTGGCTGTCCATTTCATGGAGCAGATGTCTTTGAGGACAAACTTCCCGA 8580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 8521 TAGCCGTGTGTTGGCTGTCCATTTCATGGAGCAGATGTCTTTGAGGACAAACTTCCCGA 8580

Db 8581 GACTGTGACCTTTGACTGTGATGGGTATGGGAAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 8581 GACTGTGACCTTTGACTGTGATGGGTATGGGAAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640

Db 8641 CATCATTGCTGTGTCTCTGGTATTGAGGCTTTCTCGGTGTGCGCTACACCAACGCTGA 8700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 8641 CATCATTGCTGTGTCTCTGGTATTGAGGCTTTCTCGGTGTGCGCTACACCAACGCTGA 8700

Db 8701 GATCCTCAGAGTTTCCCAATCACTAACAGACATGACCATGCCCCCCTGCGAGCCTGGCG 8760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 8701 GATCCTCAGAGTTTCCCAATCACTAACAGACATGACCATGCCCCCCTGCGAGCCTGGCG 8760

Db 8761 AAAGAAAGCCAGGGCGGTGCTCCTGCCAGCGCGTGGGCGGAGCACACGCAAAATT 8820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 8761 AAAGAAAGCCAGGGCGGTGCTCCTGCCAGCGCGTGGGCGGAGCACACGCAAAATT 8820

Db 8821 GGCTCGCTTCCTTCTCTGCATGCTACATCTAGACCCTACACTCTACACGATTTGGATAAGACGAG 8880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 8821 GGCTCGCTTCCTTCTCTGCATGCTACATCTAGACCCTACACTCTACACGATTTGGATAAGACGAG 8880

Db 8881 CGTTGGCTCGGTACACCACTTTACTATTGTGATGTTTACTCCCCGGGGATGTGTT 8940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 8881 CGTTGGCTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGGGGATGTGTT 8940

Db 8941 TATTACACCAGAAGATTGCAGAAGTTCCTTGTGAAGTATTGGCGTGTCATTGTTTT 9000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 8941 TATTACACCAGAAGATTGCAGAAGTTCCTTGTGAAGTATTGGCGTGTCATTGTTTT 9000

Db 9001 TGCCCTAGGGCTCATTGCTTGTGTTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 9001 TGCCCTAGGGCTCATTGCTTGTGTTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060

Db 9061 CTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCCGCAACAGGGGAACCCCGGGC 9120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 9061 CTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCCGCAACAGGGGAACCCCGGGC 9120

Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 6960
    AGGAGTGCATGTCCTGAAAGCCTGCAACGAAGTGAAGCCGTTGAGAAGGTCCTTCAAACCT

Qy 6901 AGGAGTGCATGTCCTGAAAGCCTGCAACGAAGTGAAGCCGTTGAGAAGGTCCTTCAAACCT 6960

Db 6961 CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGCCATGCCGATGCCCTGTTGGGAGCGGGTGA 7020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 6961 CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGCCATGCCGATGCCCTGTTGGGAGCGGGTGA 7020

Db 7021 GTGTAACCCTTTCACTGAATTGGATGTGCAATGACAGGCGAGGCCCTGATGA 7080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7021 GTGTAACCCTTTCACTGAATTGGATGTGCAATGACAGGCGAGGCCCTGATGA 7080

Db 7081 TTTACCCAGTTACCCTCCCAAAAAGGAGGTCTCTGCAGACAGGCGAAAGTGGTCGAC 7140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7081 TTTACCCAGTTACCCTCCCAAAAAGGAGGTCTCTGAATGGTCAGACGAAAGTGGTCGAC 7140

Db 7141 GGCTACAACCGCTTCCAGCTACGTTACTGGCCCCCGTACCCTTAAGATACGGGGAAAGGA 7200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7141 GGCTACAACCGCTTCCAGCTACGTTACTGGCCCCCGTACCCTTAAGATACGGGGAAAGGA 7200

Db 7201 TTCCACTCAGTCAGCCCCCGCCAAACGGCCTACAAAAAGAGAGTTGGGAAAGAGTGAGTT 7260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7201 TTCCACTCAGTCAGCCCCCGCCAAACGGCCTACAAAAAGAGAGTTGGGAAAGAGTGAGTT 7260

Db 7261 TTCGTGCAGCATGAGCTACACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7261 TTCGTGCAGCATGAGCTACACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320

Db 7321 TCTGTCTGCAACTCGGGCCATCACTAGTGGTTTCCTCAAACAAAGATCATTGGTGTATGT 7380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7321 TCTGTCTGCAACTCGGGCCATCACTAGTGGTTTCCTCAAACAAAGATCATTGGTGTATGT 7380

Db 7381 GACTGAGCGCGCGGGATGCGGAGCTTAGAAAACAAAGTCACTATTAATAGACAAACCTCT 7440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7381 GACTGAGCGCGCGGGATGCGGAGCTTAGAAAACAAAGTCACTATTAATAGACAAACCTCT 7440

Db 7441 GTTCCCCCCATCATACCACCACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7441 GTTCCCCCCATCATACCACCACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500

Db 7501 CGGTGTCATGTGGGACTATGATGCAGCTAGCAGCTCACACGCCCTCTCAAGTCTGCTAAGTC 7560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7501 CGGTGTCATGTGGGACTATGATGCAGCTAGCAGCTCACACGCCCTCTCAAGTCTGCTAAGTC 7560

Db 7561 CCACATCACTGGCCTTCGGGGCACTGATGTTCGTTCTGGAGCAGCCCGCAAGGCTGTTCT 7620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7561 CCACATCACTGGCCTTCGGGGCACTGATGTTCGTTCTGGAGCAGCCCGCAAGGCTGTTCT 7620

Db 7621 GGACTTGCAGAGAGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7621 GGACTTGCAGAGAGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680

Db 7681 AGTTCCAAAGGAGGAGTCTTCGTGAAGACAACCAACAAAGAAACCCCCAAG 7740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7681 AGTTCCAAAGGAGGAGTCTTCGTGAAGACAACCAACAAAGAAACCCCCAAG 7740

Db 7741 GCTTATCTCGTACCCCGTACACCCTCACCTTGAAATGAGATGTGTTGAGAACGATGTGTTACTACGGTCAGGT 7800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7741 GCTTATCTCGTACCCCGTACACCCTCACCTTGAAATGAGATGTGTTGAGAACGATGTGTTACTACGGTCAGGT 7800

Db 7801 TGCTCCTGACGTAGTTAAAAGCTGTCATGGGAGTGTACGCGGGTTTGTAGATCCACGTAC 7860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7801 TGCTCCTGACGTAGTTAAAAGCTGTCATGGGAGTGTACGCGGGTTTGTAGATCCACGTAC 7860

Db 7861 CCGTGTCAAGCGTCGTTGTGATGTGTCGATGTGTCGAGCCGAGCCACATGCGGATAC 7920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7861 CCGTGTCAAGCGTCGTTGTGATGTGTCGATGTGTCGAGCCGAGCCACATGCGGATAC 7920

Db 7921 AGTGTGTTTGACAGTACCATCACACACCCGAGGATATCATGGTGGAGACAGACATCTACTC 7980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7921 AGTGTGTTTGACAGTACCATCACACACCCGAGGATATCATGGTGGAGACAGACATCTACTC 7980

Db 7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCATTCACACCATTGCCGAGGCAGTTATA 8040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCATTCACACCATTGCCGAGGCAGTTATA 8040

```
Qy    9121  TTAACGACCCCGCCGATGTGAGTTTCGGCGACCATGCTGGATCAGAACCGTTTCGGCGTGAA  9180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9121  TTAACGACCCCGCCGATGTGAGTTTCGGCGACCATGCTGGATCAGAACCGTTTCGGCGTGAA  9180

Qy    9181  GCCATGTGCTGAAGGGGATGAGACGTCCCTTCTGGCTCATCCACAAAAACCGTCTCGGCTGG  9240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9181  GCCATGTGCTGAAGGGGATGAGACGTCCCTTCTGGCTCATCCACAAAAACCGTCTCGGCTGG  9240

Qy    9241  GTGAGGAGTCCTGGCTGTGTGGGAAGCAGTCAGTATAATTCCCGTCGTGTGGTGACGC  9300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9241  GTGAGGAGTCCTGGCTGTGTGGGAAGCAGTCAGTATAATTCCCGTCGTGTGGTGGTGACGC  9300

Qy    9301  CTCACGACGTATTTGTCCGCGTGTGCGAGAGCCTAGTACCCAAGGGCTGCACCCCGGTTTTG  9360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9301  CTCACGACGTATTTGTCCGCGTGTGCGAGAGCCTAGTACCCAAGGGCTGCACCCCGGTTTTG  9360

Qy    9361  TTCCAAGGCGGAGGGCAACCCCCGCCTTGGAATTAAAAACT  9399
            |||||||||||||||||||||||||||||||||||
Db    9361  TTCCAAGGCGGAGGGCAACCCCCGCCTTGGAATTAAAAACT  9399
```

FEATURES   Location/Qualifiers
     source   1..9399
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
     5'UTR   1..445
     CDS   446..9040
                     /codon_start=1
                     /transl_table=11
                     /product="polyprotein"
                     /protein_id="AAP57528.1"
                     /db_xref="GI:33090378"
                     /translation="MPVISTQTSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKV
                     QRDARPRNYKIAGIHDGLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWNIGDVT
                     THTPLVGPLVAGAVVRPVCQIVRLLEDGVNATGWFGVHLFVVCLLSLACPCSGARVT
                     DPDTNTTILTNCCQRNQVIYSPSTCLHEPGCVICADECWPANPYISHSNWTGTDS
                     FLADHIDFVMGALVTCDDALDIGELCGACVLVGDWLVRHWLIHIDLNETGTCYLLEVPTG
                     IDPGFLGFIGWMAGKVEAVIFLTKIASQVPYAIATMFSSVHYLAVGALLYYASRGKWY
                     QLLALMLYIEATSGNPIRVPTGCSIAEFCSPLMIPCPCHSYLSENVSRVICYSPKWT
                     RPITLEYNNSISWYPYTIPGARGCMVKFKNTWGCCRIRNVPSYCTMGTDAVWNDTRN
                     TYEVCGVTPWLTTAWHNGSALKLAILQYPGSKEMFKPHNWMSGHLYFEGSDTPIVYFY
                     DPVNSTLLPERWARLPGTPPVVRGSWLQVPQGFYSDVKDLATGLITKDKAWKNYQVL
                     YSATGALSLTGVTKAVVLILLGLCGSKYLILAYLCYLSLCFGRASGYPLRPVLPSQS
                     YLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAALLGFVPMAAGLPLTFVAAAAQ
                     PDYDMWRLLVAGLVLMAGRNRGHRIALLVGPMYLVALLTLLHLVTPASAFDTRIIGG
                     LTIPPVVALVVMSRGFPFAHLLPRCALVNSYLMWQRWENWFWNVTLRPBRFFLVLCFP"

```
                                          3'UTR
ORIGIN

Query Match        99.8%;  Score 9376.6;  DB 12;  Length 9399;
Best Local Similarity  99.9%;  Pred. No. 0;
Matches 9385;  Conservative 0;  Mismatches 14;  Indels 0;  Gaps  0;

Qy       1  ACCACAAACACTCCAGTTTGTTACACTCCGCTAGAATGCTCCTGGAGCACCCCCCTAG  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1  ACCACAAACACTCCAGTTTGTTACACTCCGCTAGAATGCTCCTGGAGCACCCCCCTAG  60

Qy      61  CAGGGGCGTGGGGGGATTTCCCCTGCCCGTCTGCAGAAGGGTGAGCCAACCACCTTAGTAT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61  CAGGGGCGTGGGGGGATTTCCCCTGCCCGTCTGCAGAAGGGTGAGCCAACCACCTTAGTAT  120

Qy     121  GTAGGCGGCGGGACTCATGACCGTCGCGTGCGTGGTGGTGGTGACTTGACTTGGATGGC  180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121  GTAGGCGGCGGGACTCATGACCGTCGCGTGCGTGGTGGTGGTGGTGACTTGACTTGGATGGC  180

Qy     181  CCTGATGGGCGTTCATGGGTTCGGTGCGCTTTAGGCAGCAGCTTAGGCAGCCTCCAGCCCACCA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181  CCTGATGGGCGTTCATGGGTTCGGTGCGCTTTAGGCAGCAGCTTAGGCAGCCTCCACGCCCACCA  240

Qy     241  CCTCCCAGATAGAGCGCGGCGGCACTGTAGGGAAGACCGGGGACCGGCCAAGCCACTACCAAGGACG  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241  CCTCCCAGATAGAGCGCGGCGGCACTGTAGGGAAGACCGGGGACCGGCCAAGCCACTACCAAGGACG  300

Qy     301  CAGACCTCTTTTGAGTATCACGCCTCCGGAGTAGTGGGCAAGCCCACCTATATGTGT  360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301  CAGACCCTCTTTTGAGTATCACGCCTCCGGAGTAGTGGGCAAGCCCACCTATATGT  360

Qy     361  TGGGATGGTTGGGGTTAGCATGCCTGATAGGGTCCTTGCGAGGGGAT  420
            ||||||||||||||||||||||||||||||||||||||||||||||
Db     361  TGGGATGGTTGGGGTTAGCATGCCTGATAGGGTCCTTGCGAGGGGAT  420

Qy     421  CTGGGAGTCCTGAGACCGTAGCACATGCCGTGTATTTCTACTCAACAAGTCCTGTACC  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
GATYDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVRLGKCHAWYSHYVLKFF
LLVFGENGVFFYKHLHGDVLPNDFASKLPLQEPFFPFPEGKARVRNEGRRLACGDTVD
GLPVVARLGDLVFAGLAMPPDGWAITAPFTTQCLSRGTLSMAVVMTGIDFRTWTGT
IFRLGSLATSYMGFVCDNVLYTAHIGSKGRRLAHPTGSIHPITVDAANDQIYQPPCG
AGSLTRCSCGETKGYLVTRLGSLVEVNKSDDPYWCVCGALPMAVAKGSSGAPILCSSG
HVIGMPTAARNSGGSVSQIRVRPLVCAGYHPTAHATLDTKPTVNEYSVQILIAPT
GSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCTNTGAS
LTYSTYGMYLTGACSRNYDVIICDECHATDATTVLGIGKVLTEAPSKGVRLVVLATAT
PPGVIPTPHANITEIQLTDEGTIPPFHGKKIKBENLKKGRHLIFEATKKHCDELANELA
RKGITAVSYRGCDISKIPBGDCVVVATDALCTGYTGDFDSVYDCSLMVEGTCHVDLD
PTFTMGVRVCGVSAIVKGQRRGRTGRRAGIYYVDGSCTPSGMVPECNIVEAFDAAK
AWYGLSSTEAQTILDYRTQPGLPAIGANLDEWADLFSWTNPEPSFVNTAKRTADNYV
LLTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQ
MCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSIASVPTGATVAPVVDEEE
IVEECASFIPLEAMVAAIDKLKSTITTTSPFTLETALEKLNTFLGPHAATTLAIIEYC
CGLVTLPDNFPASCVPAFIAGITTPLPHKIKMFLSLFGGAIASKLTDARGALAFMMAG
AAGTALGTWTSVGFVFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLVYSAFNPAA
GVVGVLSACAMFALTTAGPDHWPNRRLLTMLARSNTVCNEYFIATRDIRRKLIGILEAS
TPWSVISACIRWLHTPTEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGC
PFYSCQKGYKGPWIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYMRGAVPVN
ARLCGSARPDTDWTSLVVNYGVRDYCKYEKMGDHIFVTAVSSPHIFVTQVPPTLRAA
VAVDGVQVQCYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDAL
ETNDCNSTNNTPSDEAAVSALVFKQELRRTNQLLEAISAGVDTTKLPAPSIEEVVVRK
RQFRARTGSLTLPPPPRSVPGVSCPESLGRDSLPGPSNLPPSPVVLQLAMPMLLGA
GECNPFTAIGCAMTETGGGPDLPSYPPKEVSEWSDESWSTATTVSSYVTGPPYPKI
RGKDSTQSAPAKRPTKKKLGKSEFSCSMSYTWTDVISFKTASKVLSATRAITSGFLKQ
RSLVVTEPRDAELRKQKVTINRQPLFPPSYHQVRLAKEKASKVVGWVDEVAAH
TPSKSAKSHITGLRGTDVRSGAARAVLDLQKCVEAGEIPSHYRQTVIVPKEEVVKT
PQKPTKKPRRLISYPHLEMRCVEKMYYGQVAPDVVKAVMGDAYGFVDPRTRVKRLLSM
WSPDAVGATCDTVCFDSTTIPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMI
AYDGREIGYRRCRSSGVYTTSSSNSLTCWLKVNAAABQAGMKNPRFLICGDDCTVIWK
SAGADADKQAMRVFASWMKVMGAPQDCVPQPKYSLEELTSCSSNVTSGITKSGKPYYF
LTRDRPIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHPMEQMLFEDKLPE
TVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNABILRVSQSLTDWTMPPLRA
WRKKARAVLASAKRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFNYCDVYSPE
GDVPITPQRRLQKFLVKYLAVIVFALGLIAVGLAIS"
9041. .9399
```

```
Db   421  CTGGGAGTCTCGTAGACCATGCCTCTGTATTTCTACTCAAACAAGTCCTGTACC          480

Qy   481  TGCGCCCAGAACAAGCAAGACACGCAGCGTTCATATCCTGTGTCCATTAAAAAC          540
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   481  TGCGCCCAGAACAAGCAAGACGCGCAGCGTTCATATCCTGTGTCCATTAAAAAC          540

Qy   541  ATCTGTTGAAAGGGGACAACGAGCAAAGCGCAAAGTCCAGCGCGATGCTCGGCCTAA       600
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  ATCTGTTGAAAGGGGACAACGAGCAAAGCGCAAAGTCCAGCGCGATGCTCGGCCTAA       600

Qy   601  TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC    660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC    660

Qy   661  TCATGGTTGGGGACGCGCCAAGAACCCTCGGCCCATAAGTCTTCGCAATCTTCGGAATCTTCTGGA  720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  TCATGGTTGGGGACGCGCCAAGAACCCTCGGCCCATAAGTCTTCGCAATCTTCGGAATCTTCTGGA  720

Qy   721  TTACCCTTTGGGGTGGATTGGTGATGTTACACTCACACACCTTCAGTAGGCCCGCTGGT      780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  TTACCCTTTGGGGTGGATTGGTGATGTTACACTCACACACCTTCAGTAGGCCCGCTGGT      780

Qy   781  GGCAGGAGCGGGTCGTTGCGACCAGTACGCTTGCTGCTGGAGGATGGAGTCAA           840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  GGCAGGAGCGGGTCGTTGCGACCAGTACGCTTGCTGCTGGAGGATGGAGTCAA           840

Qy   841  CTGGGGCACTCGTTGGTTCGGTTCGGTGTGTATGTCTGCTATCTTTGGCGCTG           900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  CTGGGGCACTCGTTGGTTCGGTTCGGTGTGTATGTCTGCTATCTTTGGCGCTG           900

Qy   901  TCCCTGTAGTGGGGCGCGGGTCACTGACCAGACACAAATACCACAATCCTGACCAATTG     960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  TCCCTGTAGTGGGGCGCGGGTCACTGACCAGACACAAATACCACAATCCTGACCAATTG     960

Qy  961   CTGCCAGCGTAATCAGGTTATCTTATTGTCTCCTTCACTTGCCTACACGGAGCCTGGTTG   1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961   CTGCCAGCGTAATCAGGTTATCTTATTGTCTCCTTCACTTGCCTACACGGAGCCTGGTTG   1020

Qy  1021  TGTGATCTGTGCGGACGAGTGCTCGGGGTTCCCGCCAATCGGTACATCTTCACACCTTCCAA   1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  TGTGATCTGTGCGGACGAGTGCTCGGGGTTCCCGCCAATCGGTACATCTTCACACCTTCCAA   1080

Qy  1081  TTGGACTGGCACCGACTCCTTCTGGCTGACCATTGATTTTGTTATGGGGCCTCTTGT       1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  TTGGACTGGCACCGACTCCTTCTGGCTGACCATTGATTTTGTTATGGGGCCTCTTGT       1140

Qy  1141  GACCTGTGACGGCCCTTGACATTGGTAGTTGTGTGGGTGCGTGTATTAGTCGGTGACTG     1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  GACCTGTGACGGCCCTTGACATTGGTAGTTGTGTGGGTGCGTGTATTAGTCGGTGACTG     1200

Qy  1201  GCTTGTCAGGCACTGGCTTATTCACAATAGACCTCAATGAAAACTGGTACTTGTTACCTGGA   1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GCTTGTCAGGCACTGGCTTATTCACAATAGACCTCAATGAAAACTGGTACTTGTTACCTGGA   1260

Qy  1261  AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT    1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT    1320

Qy  1321  CGAGGGTATCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGGCTATTGCGACTAT   1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  CGAGGGTATCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGGCTATTGCGACTAT   1380

Qy  1381  GTTTAGCAGTGTACACTACCTGCCGGTTGGCCGTCCGATCGTTTGCTTTACATAGAAGCGACCTAT   1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  GTTTAGCAGTGTACACTACCTGCCGGTTGGCCGTCCGATCGTTTGCTTTACATAGAAGCGACCTAT   1440

Qy  1441  GTGGTATCAGTTGCTCCTCCTAGGCGTTGCTTATGTCTTTACAGAAGCGACCTCTGAAACCCCAT    1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db  1441  GTGGTATCAGTTGCTCCTCCTAGGCGTTGCTTATGTCTTTACAGAAGCGACCTCTGAAACCCCTAT    1500

Qy  1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTTGCTCGCCTTTGATGATACCATCGTCC   1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTTGCTCGCCTTTGATGATACCATCGTCC   1560
```

```
Qy  1561  TTGCCACTCTTATTTGAGTGAGAAATGTGTCAGAAGTCATTTGTTACAGTGTCCAAAGTGGAC    1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  TTGCCACTCTTATTTGAGTGAGAAATGTGTCAGAAGTCATTTGTTACAGTGTCCAAAGTGGAC    1620

Qy  1621  CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCTATACAATCCCTGAA      1680
          |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  CAGGCCTATCATCCTAGAGTATAACAACTCCATATCTTGGTACCCTATACAATCCCTGAA      1680

Qy  1681  TGCCAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA     1740
          ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  1681  TGCCAGGGGATGTATGGTTAAATTCAAAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA     1740

Qy  1741  TGTGCCATCGTACTGCACTATGGCACTGATGCAGTGTGGAACGACACTCGCAACACTTA     1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  TGTGCCATCGTACTGCACTATGGCACTGATGCAGTGTGGAACGACACTCGCAACACTTA     1800

Qy  1801  CGAAGCATGCGGTGTAACACCATGGCTAACACCATGGCAACGGCTCAGCCCTGAA        1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  CGAAGCATGCGGTGTAACACCATGGCTAACACCATGGCAACGGCTCAGCCCTGAA        1860

Qy  1861  ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC     1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC     1920

Qy  1921  AGGCCATTGTATTTTGAGGGATCAGATACCCCTATAGTTACTTTTATGACCCTGTGAA     1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  AGGCCATTGTATTTTGAGGGATCAGATACCCCTATAGTTACTTTTATGACCCTGTGAA     1980

Qy  1981  TTCCACTCTTCCACCGGAGAGGTGGCTAGGTTGCCGGTTCCGGTACCCACCTGTGGTACG   2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  TTCCACTCTTCCACCGGAGAGGTGGCTAGGTTGCCGGTTCCGGTACCCACCTGTGGTACG   2040

Qy  2041  TGGTTCTTGGTTACAGGTTACAGTTTACACGTGATGTGAAAGACTAGCCACAGG       2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  TGGTTCTTGGTTACAGGTTACAGTTTACACGTGATGTGAAAGACTAGCCACAGG       2100

Qy  2101  ATTGATCACCAAAGACAAAGCCTTGAAAAATTATCAGGTCTTATATTCCCGCCAGGTGC   2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  ATTGATCACCAAAGACAAAGCCTTGAAAAATTATCAGGTCTTATATTCCCGCCAGGTGC   2160

Qy  2161  TTTGTCTCCTTACGGGAGTTACCACCAAGGCCGTGGTGCTAATTCTGTTGGGGTTGTGTGG    2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TTTGTCTCCTTACGGGAGTTACCACCAAGGCCGTGGTGCTAATTCTGTTGGGGTTGTGTGG    2220

Qy  2221  CAGCAAGTATCTTATTTTAGCCTACTTTTAGCCCTGTCTGTCCCTTTGTTTTTGGGGCGCTTC   2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  CAGCAAGTATCTTATTTTAGCCTACTTTTAGCCCTGTCTGTCCCTTTGTTTTTGGGGCGCTTC   2280

Qy  2281  TGGTTACCCTTGCGTCTGTGTCCATCCCAGTCGTATCTCCAAGCTGGCTATCTCCG       2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TGGTTACCCTTGCGTCTGTGTCCATCCCAGTCGTATCTCCAAGCTGGCTATCTCCG       2340

Qy  2341  TTTGTCTAAAGCTCAAGTAGCTCAAGTTTCTTCATCTGTTTGCTTTGTTATCTCCG       2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  TTTGTCTAAAGCTCAAGTAGCTCAAGTTTCTTCATCTGTTTGCTTTGTTATCTCCG       2400

Qy  2401  CTGCAGGCTACGTATGCGTCGCCATGGCTTGTGCCCATGGCCGTAACCGTGCGCGGCCTTC    2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  CTGCAGGCTACGTATGCGTCGCCATGGCTTGTGCCCATGGCCGTAACCGTGCGCGGCCTTC    2460

Qy  2461  AACTTTCTTTGTTGCAGCAGCTGCTGCCAACCAGATTATGACTGGTGTGCGGACTGCT     2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  AACTTTCTTTGTTGCAGCAGCTGCTGCCAACCAGATTATGACTGGTGTGCGGACTGCT     2520

Qy  2521  AGTTGGCAGGGTTAGTTTTTGTTGCAGCAGGTTCACCGGTAACCGTGCTCGTGCTTGT      2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  AGTTGGCAGGGTTAGTTTTTGTTGCAGCAGGTTCACCGGTAACCGTGCTCGTGCTTGT      2580

Qy  2581  AGGTCCTTGGCCTTCGGTAGCGCTTTAACCCTCGGTTGGTTTACGCGGCTTGCTTCAGC    2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  AGGTCCTTGGCCTTCGGTAGCGCTTTAACCCTCGGTTGGTTTACGCGGCTTGCTTCAGC    2640
```

This page contains a two-sequence DNA alignment (Qy = query, Db = database) printed in landscape orientation. Each block shows the query sequence, a row of match bars (|||), and the database sequence, with flanking base-position numbers.

Left column (positions 2641–3780):

```
Qy  2641  TTTTGATACCGGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT  2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  TTTTGATACCGGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT  2700

Qy  2701  GTCTCGTTTGGCTTCTTTGGCTTGTTGCTTTAGTTAACTCCTATCT  2760
Db  2701  GTCTCGTTTGGCTTCTTTGGCTTGTTGCTTTAGTTAACTCCTATCT  2760

Qy  2761  TTGGCAACGTTGGGAGAATTGGTTTGGAACGTTACACTAAGACCGGAGAGTTTTTCCT  2820
Db  2761  TTGGCAACGTTGGGAGAATTGGTTTGGAACGTTACACTAAGACCGGAGAGTTTTTCCT  2820

Qy  2821  TGTGCTGGTTGTTGTTCCCCGGTGCGACATATGACGGCGCTGGTGACTTTCTGTGTGTGTCA  2880
Db  2821  TGTGCTGGTTGTTGTTCCCCGGTGCGACATATGACGGCGCTGGTGACTTTCTGTGTGTGTCA  2880

Qy  2881  CGTAGCTCTTCTATGTTAACATCCAGTGCAGCATCGTTCTTGGACTGACTCTAGGGT  2940
Db  2881  CGTAGCTCTTCTATGTTAACATCCAGTGCAGCATCGTTCTTGGACTGACTCTAGGGT  2940

Qy  2941  TAGGGCCATAGAATGTTGGTGCGTCTCGGAAAGTGTCATGCTTGTATTCTCATTATGT  3000
Db  2941  TAGGGCCATAGAATGTTGGTGCGTCTCGGAAAGTGTCATGCTTGTATTCTCATTATGT  3000

Qy  3001  TCTTAAGTTTTTCCCTCTTAGTGTTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA  3060
Db  3001  TCTTAAGTTTTTCCCTCTTAGTGTTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA  3060

Qy  3061  TGGTGATGGTCTTGCCTAATGATTTGCCTGAAACTACCATTGCAAGAGCCATTTTCCC  3120
Db  3061  TGGTGATGGTCTTGCCTAATGATTTGCCTGAAACTACCATTGCAAGAGCCATTTTCCC  3120

Qy  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAAGACGCTTGGCGTGTGTGGGGACAC  3180
Db  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAAGACGCTTGGCGTGTGTGGGGACAC  3180

Qy  3181  GGTTGATGGTTTGCCCGTTGTGCGGGGACCTTGTTTTCGCAGGGTTGGCTAT  3240
Db  3181  GGTTGATGGTTTGCCCGTTGTGCGGGGACCTTGTTTTCGCAGGGTTGGCTAT  3240

Qy  3241  GCGCGCCAGATGGGGTGGGCCATTACCGGCACCTTTACGCTGCTCTCTGAACGTGG  3300
Db  3241  GCGCGCCAGATGGGGTGGGCCATTACCGGCACCTTTACGCTGCTCTCTGAACGTGG  3300

Qy  3301  CACGCTGTCAGCGATGCAGTGGTCATGACTGTATAGACCCCCGAACTTGGACTCGAAC  3360
Db  3301  CACGCTGTCAGCGATGCAGTGGTCATGACTGTATAGACCCCCGAACTTGGACTCGAAC  3360

Qy  3361  TATCTTCAGATTAGGATCTCTGGCCACTAGCTACAAGGGGCATTGTTGTGACAACGTGTT  3420
Db  3361  TATCTTCAGATTAGGATCTCTGGCCACTAGCTACAAGGGGCATTGTTGTGACAACGTGTT  3420

Qy  3421  GTATACTGCTCACCATGGCAGCAAGGGCGCGGCTAATGACCAGGACATCTATCCACCACCATGTGGAGCTGG  3480
Db  3421  GTATACTGCTCACCATGGCAGCAAGGGCGCGGCTAATGACCAGGACATCTATCCACCACCATGTGGAGCTGG  3480

Qy  3481  CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG  3540
Db  3481  CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG  3540

Qy  3541  GTCCCTTACTCGGTGCTCTTGCGGTGTCTTGCGGGGTATCGGTGTAACACGGACTGGGGTC  3600
Db  3541  GTCCCTTACTCGGTGCTCTTGCGGTGTCTTGCGGGGTATCGGTGTAACACGGACTGGGGTC  3600

Qy  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGCGGGGCCTTCCCAT  3660
Db  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGCGGGGCCCCTTCCCAT  3660

Qy  3661  GGCTGTTGCCAAGGGTTCTTCAGGTGCGCCCCGATTCGTGCTTGTCGTGCTCCGGGCATGTTATTGG  3720
Db  3661  GGCTGTTGCCAAGGGTTCTTCAGGTGCGCCCCGATTCGTGCTTGTCGTGCTCCGGGCATGTTATTGG  3720

Qy  3721  GATGTTCACCGCTGCTAGAAATTCTGGCGGTTCAGTCAGTCAGTCAGATTAGGGTTAGGCCGTT  3780
```

Right column (positions 3721–4860):

```
Db  3721  GATGTTCACCGCTGCTAGAAATTCTGGCGGTTCAGTCAGTCAGATTAGGGTTAGGCCGTT  3780

Qy  3781  GGTGTGTGCTGATACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC  3840
Db  3781  GGTGTGTGCTGATACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC  3840

Qy  3841  TGTGCCTAACGAGTATTCAGTGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC  3900
Db  3841  TGTGCCTAACGAGTATTCAGTGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC  3900

Qy  3901  CAAATTACCACTTTCTTACATGGAGGAAGTATGAGTCTTGGTCCTAAATCCCAGTGT  3960
Db  3901  CAAATTACCACTTTCTTACATGGAGGAAGTATGAGTCTTGGTCCTAAATCCCAGTGT  3960

Qy  3961  GGCTACAACAGCATCAATGCCAAAGTACAATGCACGCGGACGTACGGCGTGAATCCAAATTG  4020
Db  3961  GGCTACAACAGCATCAATGCCAAAGTACAATGCACGCGGACGTACGGCGTGAATCCAAATTG  4020

Qy  4021  CTATTTTAATGGCAAATGTACCAACACAGAGGGGCTTCACTCACTTACGTACGCACATATGGCAT  4080
Db  4021  CTATTTTAATGGCAAATGTACCAACACAGAGGGGCTTCACTCACTTACGTACGCACATATGGCAT  4080

Qy  4081  GTACCTGACCGGAGCATGTTCCCGGAACTATGATGAATCATTGTGACAAGTGCCATGC  4140
Db  4081  GTACCTGACCGGAGCATGTTCCCGGAACTATGATGAATCATTGTGACAAGTGCCATGC  4140

Qy  4141  TACCGATGCAACCACCGTGTTGGGCATTGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA  4200
Db  4141  TACCGATGCAACCACCGTGTTGGGCATTGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA  4200

Qy  4201  TGGTTAGGCTAGTGGTTCTTGCCACGGTGAGTAATCCCCTTGAGCGTAACCGATGAAAGGCACTATCCCCTTCATGGAAAAAGAT  4260
Db  4201  TGGTTAGGCTAGTGGTTCTTGCCACGGTGAGTAATCCCCTTGAGCGTAACCGATGAAAGGCACTATCCCCTTCATGGAAAAAGAT  4260

Qy  4261  CAACATAACTGAGATTCAATTAACCGATGAAGGCACCTTATCTTTGAGGGCTACCCAAAAAACACTG  4320
Db  4261  CAACATAACTGAGATTCAATTAACCGATGAAGGCACCTTATCTTTGAGGGCTACCCAAAAAACACTG  4320

Qy  4321  TAAGGAGGAAAATCTGAAGAAAGGGAGCACCTTTGTCTCTTGAAGGAGAATAACAGCTCGTCTCTTACTATAGGGG  4380
Db  4321  TAAGGAGGAAAATCTGAAGAAAGGGAGCACCTTTGTCTCTTGAAGGAGAATAACAGCTCGTCTCTTACTATAGGGG  4380

Qy  4381  TGATGAGCTTGCTAACGAGTAGCTCAACTTGGAAC  4440
Db  4381  TGATGAGCTTGCTAACGAGTAGCTCAACTTGGAAC  4440

Qy  4441  ATGTGACATCTCAAAAATCCCTGAGGGGCACTGTGTTGTCCACTGATGGTCGTGTG  4500
Db  4441  ATGTGACATCTCAAAAATCCCTGAGGGGCACTGTGTTGTCCACTGATGGTCGTGTG  4500

Qy  4501  TACAGGTGTACACTGGACTTTGATTCCGTGTATGGCTAGAAGGGCAC  4560
Db  4501  TACAGGTGTACACTGGACTTTGATTCCGTGTATGGCTAGAAGGGCAC  4560

Qy  4561  ATGCCATGGTGACCTTGACCCACTTTCACCATGGGGTTCGTGTGCGTGTGGGTTTCAGC  4620
Db  4561  ATGCCATGGTGACCTTGACCCACTTTCACCATGGGGTTCGTGTGCGTGTGGGTTTCAGC  4620

Qy  4621  AATAGTTAAAGGCCAGCCGGCCCACAGGCCCGTGGGAGCTGGCATATACTACTA  4680
Db  4621  AATAGTTAAAGGCCAGCCGGCCCACAGGCCCGTGGGAGCTGGCATATACTACTA  4680

Qy  4681  TGTAGACGGGAGTTGTACCCCTTCGGGTATGCTCGGAATGCAACATTGTTGAAGCCTT  4740
Db  4681  TGTAGACGGGAGTTGTACCCCTTCGGGTATGCTCGGAATGCAACATTGTTGAAGCCTT  4740

Qy  4741  CGACGCAGCCAAGGCATGGTATGGTTGTCATCAACAGAAGCTCAAACTATTCTGGACAC  4800
Db  4741  CGACGCAGCCAAGGCATGGTATGGTTGTCATCAACAGAAGCTCAAACTATTCTGGACAC  4800

Qy  4801  CTATCGCACCAACCTGGGTTACCTGCGATAGGAGCAAATTTGACAGAGCAAATTTGGCGGTGGCGATCT  4860
```

```
Db  4801  CTATCGCACCCAACCTCGGGTTACCTGCCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT  4860
Qy  4861  CTTTTCTATGGTCAACCCGACCTTCATTGTTGTCAATACTGCAAAAGAACTGCTGACAA  4920
          ||||||||||||| |||||||||||||| ||||||| |||||||||||||||||||||||
Db  4861  CTTTTCTATGGTCAACCCGACCTTCATTGTTGTCAATACTGCAAAAGAACTGCTGACAA  4920

Qy  4921  TTATGTTTTGTTGACTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCC  4980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921  TTATGTTTTGTTGACTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCC  4980

Qy  4981  CAATGACGCACCACGGTGGCAGGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG  5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981  CAATGACGCACCACGGTGGCAGGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG  5040

Qy  5041  GCGCTTGGACGGGCGCTGACGCGCTGCCTCCTGCGCCCAGAGCCCAGCGGTGACCAGATACCA  5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041  GCGCTTGGACGGGCGCTGACGCGCTGCCTCCTGCGCCCAGAGCCCAGCGGTGACCAGATACCA  5100

Qy  5101  AATGTGCTTCACTGAGTCAATACTTCGGGACAGCCCACTCGCTGTTGGGCTTGGAGT  5160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101  AATGTGCTTCACTGAGTCAATACTTCGGGACAGCCCACTCGCTGTTGGGCTTGGAGT  5160

Qy  5161  GGCTATGGCTTATCTAGCCATTGACACTTTGCGCGCCACTTTGTGTGCGGCCGTTGCTGTC  5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161  GGCTATGGCTTATCTAGCCATTGACACTTTGCGCGCCACTTTGTGTGCGGCCGTTGCTGTC  5220

Qy  5221  TATTACATCAGTCCCTACCGGTTGCTACTGTCGCCCCCAGTGGTTGACGAAGAAGAAATCGT  5280
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db  5221  TATTGCATCAGTCCCTACCGGTTGCTACTGTCGCCCCCAGTGGTTGCAATCGATAAGCTGAA  5280

Qy  5281  GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGGTGCTTCACATTGACAAGCTGAA  5340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281  GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGGTGCTTCACATTGACAAGCTGAA  5340

Qy  5341  GAGTACAATCACCACAACTAGTCCTTTCACATTGGAAACCGCCCTTGAAAAACTTAACAC  5400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341  GAGTACAATCACCACAACTAGTCCTTTCACATTGGAAACCGCCCTTGAAAAACTTAACAC  5400

Qy  5401  CTTTCTTGGGGCCTCATGCAGCTACAATCCTTGCTCATCATAGAGTATTGCTGTTGGTTTTAGT  5460
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db  5401  CTTTCTTGGGGCCTCATGCAGCTACAATCCTTGCTCATCATAGAGTATTGCTGTTGGTTTTAGT  5460

Qy  5461  CACTTACCTGACAACATCCCTTTGCATCATGCGTGTTTGCTTTTCATTGCGGGTATTACTAC  5520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461  CACTTACCTGACAACATCCCTTTGCATCATGCGTGTTTGCTTTTCATTGCGGGTATTACTAC  5520

Qy  5521  CCCACTACCTCACAAGATACAAAATGTTCCGTCGTATTATTGGAGGCGCAATTGCGTCCAA  5580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521  CCCACTACCTCACAAGATACAAAATGTTCCGTCGTATTATTGGAGGCGCAATTGCGTCCAA  5580

Qy  5581  GCTTACAGACGCGTAGAGGCGCACTGGGCTTCACGATGGCCGGGGGACTGCGCGGGAACAGCTCT  5640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581  GCTTACAGACGCGTAGAGGCGCACTGGGCTTCACGATGGCCGGGGGACTGCGCGGGAACAGCTCT  5640

Qy  5641  TGGTACATGGACATCGGTGGGTTTGTCTTTGATGGGTGAGTGCTGCTAGGCGGCTATGCTGCCGCTC  5700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641  TGGTACATGGACATCGGTGGGTTTGTCTTTGATGGGTGAGTGCTGCTAGGCGGCTATGCTGCCGCTC  5700

Qy  5701  ATCCACTGCTTGCTTGACATTTTAAATGCTTAAATGCTTGACATTTTAAATGCTTGACATATGGATCAGCT  5760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701  ATCCACTGCTTGCTTGACATTTTAAATGCTTAAATGCTTGACATTTTAAATGCTTGACATATGGATCAGCT  5760

Qy  5761  TGCTGGTTTAGTCTACTCCGGTTCAATCCGGCGCCAGGAGTTGTGGGGGCTCTTGTCAGC  5820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761  TGCTGGTTTAGTCTACTCCGGTTCAATCCGGCGCCAGGAGTTGTGGGGGCTCTTGTCAGC  5820

Qy  5821  TTGTGCAATGTTTGCTTTGACAACAGCAGGGCCCAGATCACTGCCCAACAGACTTCTTAC  5880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821  TTGTGCAATGTTTGCTTTGACAACAGCAGGGCCCAGATCACTGCCCAACAGACTTCTTAC  5880

Qy  5881  TATGCTTGCTAGGAGGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGCACATCCG  5940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881  TATGCTTGCTAGGAGGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGCACATCCG  5940
```

```
Qy  5941  CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT  6000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941  CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT  6000

Qy  6001  CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGGCCTCATTGCGGCCTCATGCTCGCTCC  6060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001  CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGGCCTCATTGCGGCCTCATGCTCGCTCC  6060

Qy  6061  TTGGCAGTATGTGTGCAATTCTTTGTGATTTGCTTTAATGTCTTAAAGCTGGAGTTCA  6120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061  TTGGCAGTATGTGTGCAATTCTTTGTGATTTGCTTTAATGTCTTAAAGCTGGAGTTCA  6120

Qy  6121  GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC  6180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121  GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC  6180

Qy  6181  CTTGGATTGGATCAGGTATGCTCCAAGCACGCTGTCCATGCGGTGCTGCGAACTCATCTTTTC  6240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181  CTTGGATTGGATCAGGTATGCTCCAAGCACGCTGTCCATGCGGTGCTGCGAACTCATCTTTTC  6240

Qy  6241  TGTTGAGAATGGTTTGCAAAACTTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241  TGTTGAGAATGGTTTGCAAAACTTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300

Qy  6301  AGGGGCTGTTCCAGTCAGCTGTGTGGTCTGGTCGGCTGACCGGACCCAACTGATTG  6360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301  AGGGGCTGTTCCAGTCAGCTGTGTGGTCTGGTCGGCTGACCGGACCCAACTGATTG  6360

Qy  6361  GACTAGTCTGTCTGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA  6420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361  GACTAGTCTGTCTGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA  6420

Qy  6421  TCACATTTTGTACAGCAGTATCCTCTCCAAATGTCGTTTCACCCAGTGCCCCCAAC  6480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421  TCACATTTTGTACAGCAGTATCCTCTCCAAATGTCGTTTCACCCAGTGCCCCCAAC  6480

Qy  6481  CTTGAGAGCTGCAGTGGGCCGTGACGCGGCGGTTCAGTGTTATCTAGGTGAGCCCAA  6540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481  CTTGAGAGCTGCAGTGGGCCGTGACGCGGCGGTTCAGTGTTATCTAGGTGAGCCCAA  6540

Qy  6541  AACTCCTTGGACGACAGCATCGTCTGCTTCGGTGTTGCCTGACGGGTAAGGGTAAAACTGTTAA  6600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541  AACTCCTTGGACGACAGCATCGTCTGCTTCGGTGTTGCCTGACGGGTAAGGGTAAAACTGTTAA  6600

Qy  6601  GCTTCCCTTCCGCGGTTGACGGTCACACACCTGGTGTGCCATGCAACTTAATTGCGTGA  6660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601  GCTTCCCTTCCGCGGTTGACGGTCACACACCTGGTGTGCCATGCAACTTAATTGCGTGA  6660

Qy  6661  TGCACTTGAGACAAAATGACTGTAATTCCACAACACACTCCTAGTGATGAAGCCGCAGT  6720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661  TGCACTTGAGACAAAATGACTGTAATTCCACAACACACTCCTAGTGATGAAGCCGCAGT  6720

Qy  6721  GTCCGCTCTTGTTTTCAAACAGGAGTTGCGGGCGTACAAACCAATTGCTTGAGGCAATTTC  6780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721  GTCCGCTCTTGTTTTCAAACAGGAGTTGCGGGCGTACAAACCAATTGCTTGAGGCAATTTC  6780

Qy  6781  AGCTGGCGTTGACACCAAACTGCCAGCCCCCCCATCCGAAGAGGTAGTGGTAAGAAA  6840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781  AGCTGGCGTTGACACCAAACTGCCAGCCCCCCCATCCGAAGAGGTAGTGGTAAGAAA  6840

Qy  6841  GCGCCAGTGTTCCGGCAAGAACTGGTTCGCTTGCTTCGTTCGCCTCCCCCCCCGAGATCCGTCCC  6900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841  GCGCCAGTGTTCCGGCAAGAACTGGTTCGCTTGCTTCGTTCGCCTCCCCCCCCGAGATCCGTCCC  6900

Qy  6901  AGGAGTGTCATGTCCTGAAAGCCGCAACGAAGAGTGACCCGTTAGAGGTCTTCAAACCT  6960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901  AGGAGTGTCATGTCCTGAAAGCCGCAACGAAGAGTGACCCGTTAGAGGTCTTCAAACCT  6960

Qy  6961  CCCTCCTTCCACCACTGTTCTACAGTTGGCCATGGTGGCCATGGTGGCCCCGTGTTGGGAGCGGGGTGA  7020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961  CCCTCCTTCCACCACTGTTCTACAGTTGGCCATGGTGGCCATGGTGGCCCCGTGTTGGGAGCGGGGTGA  7020
```

```
Qy  7021 GTGTAACCCGTTTCACTGCAATGGATGTGCAATGACCGAAACAGGCGGAGGCCCTGATGA 7080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021 GTGTAACCCGTTTCACTGCAATGGATGTGCAATGACCGAAACAGGCGGAGGCCCTGATGA 7080

Qy  7081 TTTACCCAGTTACCCTCCCAAAAAGGAGGAGTCTCGGAATGGTCAGAGACGAAAGTTGGTCGAC 7140
         ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081 TTTACCCAGTTACCCTCCCAAAAAGGAGGAGTCTCGGAATGGTCAGAGACGAAAGTTGGTCAAC 7140

Qy  7141 GGCTACAACCGCTTCCAGCTTACTGGCCCCCGTACCCTAAGATACGGGGAAAGGA 7200
         ||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db  7141 GGCTACAACCGTTTCCAGCTTACTGGCCCCCGTACCCTAAGATACGGGGAAAGGA 7200

Qy  7201 TTCCACTCAGTCAGCCCGCCAAACGGCTACACGCTGATGGTGCTACAAAAAGAAGAAGTTGGGAAGAGTGAGTT 7260
         |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201 TTCCACTCAGTCAGCCCGCCAAACGGCTACACGCTGATGGTGCTACAAAAAGAAGAAGTTGGGAAGAGTGAGTT 7260

Qy  7261 TTCCGTGCAGCATGAGCTGCACACCTGGACCGACGTGATTAGCTTCAAAAACTGCTTCTCTAAAGT 7320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261 TTCGTGCAGCATGAGCTGCACACCTGGACCGACGTGATTAGCTTCAAACACTGCTTCTCTAAAGT 7320

Qy  7321 TCTGTCTGCAACTCGGGCCATCACTAGTGTGTTCCTCAAACAAAGATCATTGGTGTGTATGT 7380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7321 TCTGTCTGCAACTCGGGCCATCACTAGTGTGTTCCTCAAACAAAGATCATTGGTGTGTATGT 7380

Qy  7381 GACTGAGCCGCGGGATGCGGAGCTTAGAAAACAAAGTCACTATTAATGACAACCTCT 7440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381 GACTGAGCCGCGGGATGCGGAGCTTAGAAAACAAAGTCACTATTAATAGACAACCTCT 7440

Qy  7441 GTTCCCCCATCATCACCACCAGCAAGCAAGTTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441 GTTCCCCCATCATCACCACCAGCAAGCAAGTTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500

Qy  7501 CGGTGTCATGTGGACTATGATGAAGTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTC 7560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7501 CGGTGTCATGTGGACTATGATGAAGTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTC 7560

Qy  7561 CCACATCACTGGCCTTCGGGGCACTGATGTTCGTTCTGGAGCAGCCCGCAAGGCTGTTCT 7620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7561 CCACATCACTGGCCTTCGGGGCACTGATGTTCGTTCTGGAGCAGCCCGCAAGGCTGTTCT 7620

Qy  7621 GGACTTGCAGAAGTGTGTGAGGCAGTGAGATAGCCAGTCATTATCGGCAAAACTGTGAT 7680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7621 GGACTTGCAGAAGTGTGTGAGGCAGTGAGATAGCCAGTCATTATCGGCAAAACTGTGAT 7680

Qy  7681 AGTTCCAAAGGAGGAGGAGTCTTCGTGAAGAACAACAAAGAAACACCCCCAG 7740
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7681 AGTTCCAAAGGAGGAGGAGTCTTCGTGAAGAACAACAACAAGAAACACCCCCAG 7740

Qy  7741 GCTTATCTCGTACCCCCCACCTTGAAATGAGATGTGTTGAGAAGATGCTACTACGGGTCAGGT 7800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7741 GCTTATCTCGTACCCCCCACCTTGAAATGAGATGTGTTGAGAAGATGCTACTACGGGTCAGGT 7800

Qy  7801 TGCTCCTGACGTAGTAGTAAAGCTGTCATGGGAGATGCGTACGGGTTGAGATGTCCACGTAC 7860
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7801 TGCTCCTGACGTAGTAGTAAAGCTGTCATGGGAGATGCGTACGGGTTGTCGACCCCACGTAC 7860

Qy  7861 CCGTGTCAAGGCTTTGACAGTAAACTCACACCATCACACCAGCATTGCGAGGCAGTTATA 7920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7861 AGTGTGTTTGACAGTAAACTCAGTGACCAACACCAGCATTGCGAGGCAGTTATA 7920

Qy  7921 AGTGTGTTTGACAGTAAACTCAGTGACCAACACCATCACACCGAGGAGGACATCTCACTC 7980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7921 AGCAGTAAACTCAGTGACCAACACCATCACACCGAGGAGGACATCTCACTC 7980

Qy  7981 CCGGTCAAGGCTTATCGTCGTGTGTCATGTGTCGATATCGCGAGATCGCCATGTAGGTC 8040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7981 CCGGTCAAGGCTCGTCGTGTGTCATGTGTCGATATCGCGAGATCGCCATGTAGGTC 8040

Qy  8041 CGCTGAGGACCGATGATCGCTTATGATGGCGAGATCTCGTAGTGTAGGTC 8100
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  8041 CGCTGAGGACCGATGATCGCTTATGATGGCGAGATCTCGTAGTGTAGGTC 8100

Qy  8101 TTCCGGGCGTCTATACTACCTCAAGTTCCAACAGTTTGACCTGCTGCTGAAGGTAAATGC 8160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8101 TTCCGGGCGTCTATACTACCTCAAGTTCCAACAGTTTGACCTGCTGCTGAAGGTAAATGC 8160
```

```
Qy  8161 TGCAGCCGAACAGCGGCTGGCATGAAGCACCCTCGCTTCCTTATTGCGGCGATGATTGCAC 8220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161 TGCAGCCGAACAGCGGCTGGCATGAAGCACCCTCGCTTCCTTATTGCGGCGATGATTGCAC 8220

Qy  8221 CGTAATTTGGAAGAGAGCCGGCGGAGCAGATGCAGACAAACAAGCAATGCGTGTGTCTTCGTAG 8280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8221 CGTAATTTGGAAGAGAGCCGGCGGAGCAGATGCAGACAAACAAGCAATGCGTGTGTCTTCGTAG 8280

Qy  8281 CTGGATGAAGGTGATGGGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281 CTGGATGAAGGTGATGGGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340

Qy  8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAGTGGCAAGCCTTA 8400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAGTGGCAAGCCTTA 8400

Qy  8401 CTACTTCTTACAAGATCCTCGTATCCCCCTGGCAGGTGCTCGCCGAGGGTCTCGGG 8460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8401 CTACTTCTTACAAGATCCTCGTATCCCCCTGGCAGGTGCTCGCCGAGGGTCTCGGG 8460

Qy  8461 ATACAACCCCAGTGCTGCGGTGATTGGGTATCTAATAATACATCACTACTCCATGTTGTGGGT 8520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8461 ATACAACCCCAGTGCTGCGGTGATTGGGTATCTAATAATACATCACTACTCCATGTTGTGGGT 8520

Qy  8521 TAGCCGTGTGTTGGCTGTCCATTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8521 TAGCCGTGTGTTGGCTGTCCATTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580

Qy  8581 GACTGTGACCTTTGACTGTGTATGGTATGGGATGGGAAAAATTATACGGTGGCGTGTAGCGCCAG 8640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8581 GACGGTGACCTTTGACTGTGTATGGTATGGGATGGGAAAAATTATACGGTGGCGTGTAGCGCCAG 8640

Qy  8641 CATCATTGCTGGTGTTCGTTCTCGGTGTGCGGTGTATTGAGGCTTCGGTGTGCGGTTGAGGCGTGA 8700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8641 CATCATTGCTGGTGTTCGTTCTCGGTGTGCGGTGTATTGAGGCTTCGGTGTGCGGTTGAGGCGTGA 8700

Qy  8701 GATCCTCAGAGTTCCCAATCACTACAGCGACATGACAAGTGATCCCGGCGGCCTGCGCG 8760
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8701 GATCCTCAGAGTTCCCAATCACTACAACGACATGACAAGTGATCCCGGCGGCCTGGGCG 8760

Qy  8761 AAAGAAAGCCAGGCGGTCCTCCTGGCATGCCCAAGAAACCCCAAAATT 8820
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  8761 AAAGAAAGCCAGGCGGTCCTCCTGGCATGCCCAAGAAACCCCAAAATT 8820

Qy  8821 GGCTCGCTCCTTCTCTGGCATGTCACATCTAGACCTCTACACATTCTAGACCTCAACGCTGA 8880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8821 GGCTCGCTCCTTCTCTGGCATGTCACATCTAGACCTCTACACATTCTAGACCTCAACGCTGA 8880

Qy  8881 CGTGGCTCGGTACACCACCTTTCAATTATTGTGATGTTTACTCCCCGGGGATGTGTT 8940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8881 CGTGGCTCGGTACACCACCTTTCAATTATTGTGATGTTTACTCCCCGGGGATGTGTT 8940

Qy  8941 TATTACACCACAGAGAAGATTGCAGAAGTTCCTTGTGAAGTTAGCCATCAATTGTTTT 9000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8941 TATTACACCACAGAGAAGATTGCAGAAGTTCCTTGTGAAGTTAGCCATCAATTGTTTT 9000

Qy  9001 TGCCCTAGGGCTCATTGCTGTGTGGATGGGCAACCGGCAACAGGGAGACCCCGGGC 9060
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9001 CTAACAGTTTTTTTTTTTTTTTTTTAGGGCAGCAGGGCAACAGGGAGACCCCGGGC 9060

Qy  9061 CCGGTCAAGGCTAAACTCAGTTGACCAACACCATCACACCGAGGAGGACATCTCACTC 9120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9061 TTAACGACCCCGATGTGAGTGAGTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAA 9120

Qy  9121 CGCTGAGGACCGATGATCGCTTATGATGGCGAGATCTCGTAGTGTAGGTC 9180
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9121 TTAACGACCCCGATGTGAGTGAGTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAA 9180

Qy  9181 GCCATGGTGCTGAAGGGATGAGCTCCCTCTGCGCTCATCACAAAAAACCGTCTCGGGTGG 9240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9181 GCCATGGTGCTGAAGGGATGAGCTCCCTCTGCGCTCATCACAAAAAACCGTCTCGGGTGG 9240
```

RESULT 4
AX805212
LOCUS   AX805212   9397 bp   DNA   linear   PAT 25-NOV-2003
DEFINITION Sequence 2 from Patent WO03059944.
ACCESSION   AX805212
VERSION   AX805212.1 GI:38522352
KEYWORDS   .
SOURCE   synthetic construct
ORGANISM   synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
AUTHORS   de Tomassi,A., Graziani,R., Paonessa,G. and Traboni,C.
TITLE   GB virus B based replicons and replicon enhanced cells
JOURNAL   Patent: WO 03059944-A 2 24-JUL-2003;
            ISTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGELETTI S.P.A.
            (IT)
FEATURES   Location/Qualifiers
            source   1..9397
                            /organism="synthetic construct"
                            /mol_type="unassigned DNA"
                            /db_xref="taxon:32630"
                            /note="GBV-B Replicon"
ORIGIN

Query Match   99.4%; Score 9338.6; DB 6; Length 9397;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 9368; Conservative 0; Mismatches 29; Indels 2; Gaps 1;

[Nucleotide sequence alignment — Query (Qy) versus Database (Db) — reproduced below as paired rows with their position numbers]

Qy   1 ACCACAAACACTCCAGTTGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACTCCCCCTAG 60
Db   1 ACCACAAACACTCCAGTTGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG 60

Qy   61 CAGGGGCGTGGGGGATTTCCCCTGCCCCGTCTCGCAGAGAGGGTGGAGCCAACCACTTAGTAT 120
Db   61 CAGGGGCGTGGGGGATTTCCCCTGCCCCGTCTCGCAGAGAGGGTGGAGCCAACCACTTAGTAT 120

Qy 121 GTAGGCGGCGGGGGACTCATGAGCGTCGGCGTGATGACGGCCAAGCGCCAAGCTTGACTTGGATGGC 180
Db 121 GTAGGCGGCGGGGGACTCATGAGCGTCGGCGTGATGACGGCCAAGCGCCAAGCTTGACTTGGATGGC 180

Qy 181 CCTGATGGGCCGTTCGGTGGTGGTGTCGGCGCTTAGGCAGCCCTTCCAGCGCCTCCAGGCCACCA 240
Db 181 CCTGATGGGCCGTTCGGTGGTGGTGTCGGCGCTTAGGCAGCCCTTCCAGCGCCTCCAGGCCACCA 240

Qy 241 CCTCCCAGATAGAGCGGCGGGCACTGTAGGGAAGACCCGGGAACCCGGGACACCAAGGACG 300
Db 241 CCTCCCAGATAGAGCGGCGGGCACTGTAGGGAAGACCCGGGAACCCGGGACACCAAGGACG 300

Qy 301 CAGAGCCTCTTTTGAGTATCACGCGCTACCCTCCGGAAGTGGGCAAGCCCACTTATATGT 360
Db 301 CAGACCCTCTTTTGAGTATCACGCGCTACACCTCCGGAAGTGGGCAAGCCCACTTATATGT 360

Qy 361 TGGGATGGTTGGGGGTTAGCCATCCATCCATCCGCCTGATAGGGTCCTTGCCGAGGGGAT 420
Db 361 TGGGATGGTTGGGGGTTAGCCATCCATCCGCCTGATAGGGTCCTTGCCGAGGGGAT 420

Qy 421 CTGGGAGTCTCGTAGACCGTGTATTTTCTACTCAAACAAGTCCTGTACC 480
Db   ... (alignment continues)

[Upper-right column of this page continues the alignment with position numbers 421–1560 (Qy) aligning to 9181–9399 (Db):]

Qb 9181 GCCATGGTCTGAAGGGGATGACCTCCCTTCTGCTCATCCACAAAAACCGCTCGGGTGG 9240
Qy 9241 GTGAGGAGTCCTGGCTGTGTGGGGAAGCAGTCAGTAGTATATATTTTCCCGTCGTGTGTGGTGACGC 9300
Db 9241 GTGAGGAGTCCTGGCTGTGTGGGGAAGCAGTCAGTAGTATATATTTTCCCGTCGTGTGTGGTGACGC 9300
Qy 9301 CTCACGACGTATTTGTCCGCTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTG 9360
Db 9301 CTCACGACGTATTTGTCCGCTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTG 9360
Qy 9361 TTCCAAGCGGAGGGCAACCCCCGCTTGGAATTAAAAACT 9399
Db 9361 TTCCAAGCGGAGGGCAACCCCCGCTTGGAATTAAAAACT 9399

```
Qy  1561  TTGCCACTCTTATTTGAGTGAGAATGTCTCAGAAGTCATTGTTACAGTCCAAAGTGGAC  1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  TTGCCACTCTTATTTGAGTGAGAATGTCTCAGAAGTCATTGTTACAGTCCAAAGTGGAC  1620

Qy  1621  CAGGCCCATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG  1680
          |||||||    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  CAGGCCCTGTCACTCTCAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG  1680

Qy  1681  TGCGAGGGGATGTATGGTTAAATTCAAAATAACACATGGGGTTGCTGCCGTATTCGCAA  1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  TGCGAGGGGATGTATGGTTAAATTCAAAATAACACATGGGGTTGCTGCCGTATTCGCAA  1740

Qy  1741  TGTGCCATGCTACTGCACTATGGGCACTGATGCAGTGTGGAACGACACTGCAACACTTA  1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  TGTGCCATGCTACTGCACTATGGGCACTGATGCAGTGTGGAACGACACTGCAACACTTA  1800

Qy  1801  CGAAGCATGCCGTGTAACAACCGGTGTAACAACCGGCACAACGGCTCAGCCCTGAA  1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  CGAAGCATGCCGTGTAACAACCGGTGTAACAACCGGCACAACGGCTCAGCCCTGAA  1860

Qy  1861  ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC  1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC  1920

Qy  1921  AGGCCATTGTATTTGAGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAA  1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  AGGCCATTGTATTTGAGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAA  1980

Qy  1981  TTCCACTCTCCTACCACCGGAGAGGTTGGGCTAGGTTGCCCCGTACCCCTGTGTACG  2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  TTCCACTCTCCTACCACCGGAGAGGTTGGGCTAGGTTGCCCCGTACCCCTGTGTACG  2040

Qy  2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG  2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG  2100

Qy  2101  ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTATATTCCCGCCACGGGTGC  2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTATATTCCCGCCACGGGTGC  2160

Qy  2161  TTTGTCTTACGGGAGTTACCACCCAAGGCCCGTGCTAATTCTGTTGGGGTTGTGTGG  2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TTTGTCTTACGGGAGTTACCACCCAAGGCCCGTGCTAATTCTGTTGGGGTTGTGTGG  2220

Qy  2221  CAGCAAGTATCTTATTTTAGCCTACCTCTGTACTTGTCCCTTGTTTGTTGGGCGCGCTTC  2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  CAGCAAGTATCTTATTTTAGCCTACCTCTGTACTTGTCCCTTGTTTGTTGGGCGCGCTTC  2280

Qy  2281  TGGTTACCCTTTGCGTCCTGTGCTCCCATCCCAGTCGTATCGCCAAGCTGGCTGGGATGT  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TGGTTACCCTTTGCGTCCTGTGCTCCCATCCCAGTCGTATCGCCAAGCTGGCTGGGATGT  2340

Qy  2341  TTTTGTCTAAAGCTCAAGTAGCTCAAGATTTTCTTGATTTTCTTCATCGTTGCTCATCTCCG  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  TTTTGTCTAAAGCTCAAGTAGCTCAAGATTTTCTTGATTTTCTTCATCGTTGCTCATCTCCG  2400

Qy  2401  CTGCAGGCTACGTTATGCTGCCCTTTTTAGGGTTGTGCCATGGCGTGCTGGCTGGGATGT  2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  CTGCAGGCTACGTTATGCTGCCCTTTTTAGGGTTGTGCCATGGCGTGCTGGCTGGGATGT  2460

Qy  2461  AACTTTCTTTGTTGCAGCAGGATTATGACTGGCGTCACCGGCTACCCGGTAACCTCTGCT  2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  AACTTTCTTTGTTGCAGCAGGATTATGACTGGCGTCACCGGTTACCCGGTAACCTCTGCT  2520

Qy  2521  AGTGGCAGGGTTAGTTTTGTGGGGCCGGCCGGTAACCGGTGGTCACCGCATAGTCTGTTGT  2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  AGTGGCAGGGTTAGTTTTGTGGGGCCGGCCGGTAACCGGTGGTCACCGCATAGTCTGTTGT  2580

Qy  2581  AGGTCCTTGGCCTCTGTGTAGGCGCTTTAAACCCTTCGGTTGGTTACGCCTGCTTCAGC  2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  AGGTCCTTGGCCTCTGTGTAGGCGCTTTAAACCCTTCGGTTGGTTACGCCTGCTTCAGC  2640
```

```
Qy  2641  TTTTGATACCGAGATAATTGGAGGGCTGACATATACCACCTGTGACTAGTAGCATTAGTTGTGTCAT  2700
          ||||||     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  TTTTGACACCGAGATAATTGGAGGGCTGACAATATACCACCTGTGACTAGTAGCATTAGTTGTGTCA  2700

Qy  2701  GTCTCGTTTTGGCTTCTTGCTCACTCACTTGCTTAGTTACCTCGCTGTGCTTAGTTAACTCCTATCT  2760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  GTCTCGTTTTGGCTTCTTGCTCACTCACTTGCTTAGTTACCTCGCTGTGCTTAGTTAACTCCTATCT  2760

Qy  2761  TTGGCAACGTTGGGAGAATTGGTTTTTGAACGTTACACTAAGACCGGAGAGGTTTTCCT  2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  TTGGCAACGTTGGGAGAATTGGTTTTTGAACGTTACACTAAGACCGGAGAGGTTTTTCT  2820

Qy  2821  TGTGCTGGTTGTTGTTTCCCGGTGCGACATATGACGGCGTCGGTGACTTTCTGTGTGTCA  2880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  TGTGCTGGTTGTTGTTTCCCGGTGCGACATATGACGGCGTCGGTGACTTTCTGTGTGTCA  2880

Qy  2881  CGTAGCTCTTCTATGTTAACATCCAGTCGCAGCATGTCGTTCTTTGGGACTGACTCTAGGGT  2940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  CGTAGCTCTCCTATGTTAACATCCAGTCGCAGCATGTCGTTCTTTGGGACTGACTCTAGGGT  2940

Qy  2941  TAGGGCCCATAGAATGTTGGTGCGTCTCGGAAAGTGTCATGCTGGTATTCTCATTATGT  3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941  TAGGGCCCATAGAATGTTGGTGCGTCTCGGAAAGTGTCATGCTGGTATTCTCATTATGT  3000

Qy  3001  TCTTAAGTTTTTCCTCTTAGTGTTTGGTGAGAATGTGTGTTTTCTATAAGCACTTGCA  3060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  TCTTAAGTTTTTCCTCTTAGTGTTTGGTGAGAATGTGTGTTTTCTATAAGCACTTGCA  3060

Qy  3061  TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCCATTTTCCC  3120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061  TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCCATTTTCCC  3120

Qy  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGAACGCTTGGGCGTGTGGGACAC  3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGAACGCTTGGGCGTGTGGGGACAC  3180

Qy  3181  GGTTGATGGTTGCCCGGTTGTCCCGGTGTTGCCCGACCTTGTTTTCGCAGGGGTTGGCTAT  3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181  GGTTGATGGTTGCCCGGTTGTCCCGGTGTTGCCCGACCTTGTTTTCGCAGGGGTTAGCTAT  3240

Qy  3241  GCCGCCAGATGGGGTGGGCCATTACCGCACCTTTACGCTGCAGTGTCTCTGAACGTGG  3300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241  GCCGCCAGATGGGGTGGGCCATTACCGCACCTTTACGCTGCAGTGTCTCTGAACGTGG  3300

Qy  3301  CACGCTGTCAGCGATGGCAGTGTCATGATCGTATAGACCCCGAACTTGGACTGGAAC  3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301  CACGCTGTCAGCGATGGCAGTGTCATGATCGTATAGACCCCGAACTTGGACTGGAAC  3360

Qy  3361  TATCTTCAGATTAGGATCTCTGGCCACTAGCTACATGGGATTTGTTGTGACAACGTGTT  3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361  TATCTTCAGATTAGGATCTCTGGCCACTAGCTACATGGGATTTGTTGTGACAACGTGTT  3420

Qy  3421  GTATACTGCTCACCATGGCAGCAGGGGCGCGGTTGGCTCACATCATCAACCACCATGGGTC  3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421  GTATACTGCTCACCATGGCAGCAGGGGCGCGGTTGGCTCACATCATCAACCACCATGGGTC  3480

Qy  3481  CCCAATAACGTTGACGCGGCTAATGACCGGGGAGACCAAGGGGTATCTGGAGCTGGGGTC  3540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481  CCCAATAACGTTGACGCGGCTAATGACCGGGGAGACCAAGGGGTATCTGGAGCTGGGGTC  3540

Qy  3541  GTCCCTTACTCGGTCTTGGTGTGTATCGGTGTGTGCCGGGCCCTTCCCAT  3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541  GTCCCTTACTCGGTCTTGGTGTGTATCGGTGTGTGCCGGGCCCTTCCCAT  3600

Qy  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCCTTATTGGTGTGGCCGGGCCCATTGTATTGG  3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601  ATTGGTTGAGGTCAATAAATCCGATGACCCTTATTGGTGTGGCCGGGCCCATTGTATTGG  3660

Qy  3661  GGCTGTTGCCAAGGGTTCTTCAGGGTCTTCAGGTTACGCGGTTATTGG  3720
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661  GGCTGTTGCCAAGGGTTCTTCAGGGTCTTCAGGTTACGCGGTTATTGG  3720

Qy  3721  GATGTTCACCGCTGCTAGAAATTCTGCGCGGTTCAGTCAGTGAAAATTCTGGCGGTT  3780
```

```
Qy  5941  CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT  6000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941  CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT  6000

Qy  6001  CCGTTGGCTCCACACCCCGACGGAGATGATTGCGGCCTCATTGCTTGGGGTCTAGAGAT   6060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001  CCGTTGGCTCCACACCCCGACGGAGATGATTGCGGCCTCATTGCTTGGGGTCTAGAGAT   6060

Qy  6061  TTGGCAGTATGTGTGCAATTTCTTTGTGATTGCTTAATGTCCTTAAAGCTGGAGTTCA    6120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061  TTGGCAGTACGTGTGCAATTTCTTTGTGATTGCTTAATGTCCTTAAAGCTGGAGTTCA    6120

Qy  6121  GAGCATGGTTAACATTCCTGGTGTCCTTTCTACAGCTGCCAGAGAGGGGTACAAGGGCCC  6180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121  GAGCATGGTTAACATTCCTGGTGTCCTTTCTACAGCTGCCAGAGAGGGGTACAAGGGCCC  6180

Qy  6181  CTGGATTGGATCAGGTATGCTCCAAGCACGCGTGTCCATGCCGGTGCTGAACTCATCTTTC  6240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181  CTGGATTGGATCAGGTATGCTCCAAGCACGCGTGTCCATGCCGGTGCTGAACTCATCTTTC  6240

Qy  6241  TGTTGAGAAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG    6300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241  TGTTGAGAAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG    6300

Qy  6301  AGGGGCGTGTTCCAGTCAACGCTAGGCTGTGTGTGGGTCGGCTAGACCGGACCCAACTGATTG  6360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301  AGGGGCGTGTTCCAGTCAACGCTAGGCTGTGTGTGGGTCGGCTAGACCGGACCCAACTGATTG  6360

Qy  6361  GACTAGTTCTTGTCGTCAATTATGGGCGTTAGGCTACTGTAAATATGAGAAATGGGAGA   6420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361  GACTAGTTCTTGTCGTCAATTATGGGCGTTAGGCTACTGTAAATATGAGAAATTGGGAGA  6420

Qy  6421  TCACATTTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGTGCCCCCAAC    6480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421  TCACATCTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGTGCCCCCAAC    6480

Qy  6481  CTTGAGAGCTGCAGTGCCGTGACGCGTACAGGTTCAGTTGTTATCTGAGTGTGTTATCTGAGCCCCAA  6540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481  CTTGAGAGCTGCAGTGCCGTGACGCGTACAGGTTCAGTTGTTATCTGAGTGTGTTATCTGAGCCCCAA  6540

Qy  6541  AACTTCCTTGGACACATCTGCTTGCTGCTTGACACGTAAGGGTAAAACTGTTAA       6600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541  AACTTCCTTGGACGACATCTGCTTGCTGCTTGACACGTAAGGGTAAAACTGTTAA       6600

Qy  6601  GCTTCCCTTCCGCGGTTGACGGTCACACACCTGGTGCGCAGCACCCCTGGTGTGCGCAACTTAATTTGCGTGA  6660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601  GCTTCCCTTCCGCGGTTGACGGTCACACACCTGGTGCGCAGCACCCCTGGTGTGCGCAACTTAATTTGCGTGA  6660

Qy  6661  TGCACTTGAGACAAATGACTTGTAATTCCACAACAACACTCCTAGTGATGAGCCCGAGT  6720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661  TGCACTTGAGACAAATGACTTGTAATTCCACAACAACACTCCTAGTGATGAGCCGCAGT  6720

Qy  6721  GTCGGCTCTTGTTTCAAACACCACCAATTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC  6780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721  GTCGGCTCTTGTTTCAAACACCACCAATTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC  6780

Qy  6781  AGGAGTGGCGGTTGACACCAAACTGCCCAGCCCCCCTGAAGAGGTAGTGGTAAGAAA    6840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781  AGGAGTGGCGGTTGACACCAAACTGCCCAGCCCCCCTGAAGAGGTAGTGGTAAGAAA    6840

Qy  6841  GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCCCTCCGAGATCCGTCCC  6900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841  GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCCCCCCCTCCGAGATCCGTCCC  6900

Qy  6901  AGGAGTGTCATGTCCTGAAAGCTGCGACGAGTGACACGGTTAGAAGGTCCTTCAAACCT   6960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901  AGGAGTGTCATGTCCTGAAAGCTGCGACGAGTGACACGGTTAGAAGGTCCTTCAAACCT   6960

Qy  6961  CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGCCGGATGCCCCTGTTGGGAGCGGGTGA  7020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961  CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGCCGGATGCCCCTGTTGGGAGCGGGTGA  7020
```

```
Qy  7021  GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAATGACCGAAACAGGCGGAGGGCCCTGATGA  7080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021  GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAATGACCGAAACAGGCGGAGGGCCCTGATGA  7080

Qy  7081  TTTACCCAGTTACCCTCCCAAAAGAGGAGTCTCTGAATGGTCAGCAGAAAGTTGGTCGAC  7140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081  TTTACCCAGTTACCCTCCCAAAAGAGGAGTCTCTGAATGGTCAGCAGAAAGTTGGTCAAC  7140

Qy  7141  GGCTACAACGCGTTCAGCTACTGGCCCCCCGTACCCTAAGATACGGGGAAAGA       7200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141  GGCTACAACGCGTTCAGCTACTGGCCCCCCGTACCCTAAGATACGGGGAAAGGA      7200

Qy  7201  TTCCACTCAGTCAGCCCGCCAAACGGCCTACAAAAAGAAGTTGGGAAAGAGTGAGTT  7260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201  TTCCACTCAGTCAGCCCGCCAAACGGCCTACAAAAAGAAGTTGGGAAAGAGTGAGTT  7260

Qy  7261  TTCGTGCAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTCTAAAGT  7320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261  TTCGTGCAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT  7320

Qy  7321  TCTGTCTGCAACTCGGGCCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGT  7380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7321  TCTGTCTGCAACTCGGGCCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGT  7380

Qy  7381  GACTGAGCCGCGGGATGCGGAGCTTGAGAAAACAAAAGTCACTATTAATAGACAACCTCT  7440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381  GACTGAGCCGCGGGATGCGGAGCTTGAGAAAACAAAAGTCACTATTAATAGACAACCTCT  7440

Qy  7441  GTTCCCCCATCATACCACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT  7500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441  GTTCCCCCATCATACCACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT  7500

Qy  7501  CGGTGTCATGTGGACTATGATGAAGTAGCAGTCACACGCCCCTCTAAGTCTGTAAGTC  7560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7501  CGGTGTCATGTGGACTATGATGAAGTAGCAGTCACACGCCCCTCTAAGTCTGTAAGTC  7560

Qy  7561  CCACATCACTGGCCTTCGGGGGACTGATGTTCGTTCTGGGAGCAGCCCGCAAGGCTGTTCT  7620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7561  CCACATCACTGGCCTTCGGGGGACTGATGTTCGTTCTGGGAGCAGCCCCGCAAGGCTGTTCT  7620

Qy  7621  GGACTTGCAGAGTGTGTGAGGCAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT  7680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7621  GGACTTGCAGAGTGTGTGAGGCAGGCAGGTGAGATACCGAGTCATTATCGGCAAACAGTGAT  7680

Qy  7681  AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAAACAACAAAGAAAACCCCCAAG    7740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7681  AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAAACAACAAAGAAAACCCCCAAG    7740

Qy  7741  GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGT  7800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7741  GCTCATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGT  7800

Qy  7801  TGCTCCTGACGTAGTAGTTAAAGCTGTCATGGGGATGTTGTAGATCCAGTAC       7860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7801  TGCTCCTGACGTAGTAGTTAAAGCTGTCATGGGGATGTTGTAGATCCAGTAC       7860

Qy  7861  CCGTGTCAAGCGTCTGTGTTGTCGATGTGTCACCCGATGCAGTCGGAGCCACATGCGATAC  7920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7861  CCGTGTCAAGCGTCTGTGTTGTCGATGTGTCACCCGATGCAGTCGGAGCCACATGCGATAC  7920

Qy  7921  AGTGTGTTTGACAGTACCATCACACCCGAGATATCATGGTGGAGACAGACATCTACTC  7980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7921  AGTGTGTTTGACAGTACCATCACACCCGAGATATCATGGTGGAGACAGACATCTACTC  7980

Qy  7981  AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCCATTCACACCGAGGCAGTTATA    8040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7981  AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCCATTCACACCGAGGCAGTTATA    8040

Qy  8041  CCGTGGAGGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTAGGTC  8100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8041  CCGTGGAGGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTAGGTC  8100

Qy  8101  TTCCGGCGGTCTATACTACCTCAAGTTCCAACAGTTTGACCTGCTGCTTGACCTGCTGAAGGTAAATGC  8160
```

```
Db  8101  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TTCCGGGTCTATATCACTACCTACCTGGCTGGCTGAGAAGGTAAATGC  8160

Qy  8161  TGCAGCCGAACAGGCTGCATGAGAGAACCCTGCTTCCTTATTTGCGGCGATGATTGCAC  8220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161  TGCAGCCGAACAGGCTGCATGAGAGAACCCTGCTTCCTTATTTGCGGCGATGATTGCAC  8220

Qy  8221  CGTAATTGGAAGAGCGCCGGAGCAGATGCAGACAAAGCAAGCAATGCGTGTCTTTGCTAG  8280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8221  CGTAATTGGAAAAGCGCCGGAGCAGATGCAGACAAAGCAAGCAATGCGTGTCTTTGCTAG  8280

Qy  8281  CTGGATGAAGGTGATGGGTGCACCACAGATTGTGTGCCTCAACCCAAATACAGTTTTGGA  8340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281  CTGGATGAAGGTGATGGGTGCACCACAGATTGTGTGCCTCAACCCAAATACAGTTTTGGA  8340

Qy  8341  AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCCAAAAGTGGCAAGCCTTA  8400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8341  AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCCAAAAGTGGCAAGCCTTA  8400

Qy  8401  CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTCGCCGAGGGTCTGGG  8460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8401  CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTCGCCGAGGGTCTCGG  8460

Qy  8461  ATACAACCCCAGTGCTGCGTGGATTGGGTATCTAATACATCACTACCCATGTTTGTGTGGGT  8520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8461  ATACAACCCCAGTGCTGCGTGGATTGGGTATCTAATACATCACTACCCATGTTTGTGTGGGT  8520

Qy  8521  TAGCCGTGTGTTGGGCTGTCATTTCATGAGCAGCAGATGCGTCTTTGAGGACAAACTTCCCGA  8580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8521  TAGCCGTGTGTTGGGCTGTTCCATTTCATGAGCAGCAGATGCGTCTTTGAGGACAAACTTCCCGA  8580

Qy  8581  GACTGTGACCTTTGACTGTTGGTATGGGAAAAATTATACGGTCGTGTAGAAGATCTGCCCAG  8640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8581  GACTGTGACCTTTGACTGTTGGTATGGGAAAAATTATACGGTCGTGTAGAAGATCTGCCCAG  8640

Qy  8641  CATCATTGCTGGTGTGCACGGTATTGAGGCTTTCTCGGTGGTGCGCTACACCAACGCTGA  8700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DDb 8641  CATCATTGCTGGTGTGCACGGTATTGAGGCTTTCTCGGTGGTGCGCTACACCAACGCTGA  8700

Qy  8701  GATCCTCAGAGTTCCCAATCACTAACAGACATGACCATGCCCCCCTGCGAGCCTGGCG  8760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8701  GATCCTCAGAGTTCCCAATCACTAACAGACATGACCATGCCCCCCTGCGAGCCTGGCG  8760

Qy  8761  AAAGAAGCCAGGGCCGGTTCCTCGCCAGCGCCAAGAGGCGTGGCGGAGCACACGCCAAATT  8820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8761  AAAGAAGCCAGGGCCGGTTCCTCGCCAGCGCCAAGAGGCGTGGCGGAGCACACGCCAAATT  8820

Qy  8821  GGCTCGCTCCTTCTCTGGCATGCCTACATCTACACTCTAGACCTCTACACTGGGATAAGACGAG  8880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DDb 8821  GGCTCGCTCCTTCTCTGGCATGCCTACATCTACACTCTAGACCTCTACACTGGGATAAGACGAG  8880

Qy  8881  CGTGGCTTCGCGTACACACTTCAATTATTGTGATGTTTACTCCCCGGAGGGGGGATGTGTT  8940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8881  CGTGGCTTCGCGTACACACTTCAATTATTGTGATGTTTACTCCCCGGAGGGGGGATGTGTT  8940

Qy  8941  TATTACCACCACAGAGAAGAATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTTT  9000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8941  TGTTACCACCACAGAGAAGAATTGCAGAAGTTTCTTGTGAAGTATTTGGCTGTCATTGTTTT  9000

Qy  9001  TGCCCTAGGGCTCATTGCTGTGTTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA  9060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9001  TGCCCTAGGGCTCATTGCTGTGTTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA  9060

Qy  9061  CTAACAGTTTTTTTTTTTTTTTTTTAGGGCAGCAGCGGAGACAGGGGAGACCCCCGGGC  9120
          ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9061  CTAACAG--TTTTTTTTTTTTTTTTTTAGGGCAGCAGCGGAGACAGGGGAGACCCCCGGGC  9118

Qy  9121  TTAACGCACCCCGCCATGGTGAGTTTGGCGACCATGGTGATCAGAACCGTTTCGGGGTGAA  9180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9119  TTAACGCACCCCGCCATGGTGAGTTTGGCGACCATGGTGATCAGAACCGTTTCGGGGTGAA  9178

Qy  9181  GCCATGGTGCTGAAGGGGATGAGCTCCCTTCTGGCTCATCCACAAAAACCGTCTCGGGTGG  9240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  9179  GCCATGGTCTGAAGGGGATGAGCTCCCTTCTGGCTCATCCACAAAAACCGTCTCGGGTGG  9238

Qy  9241  GTGAGGAGTCCTGGCGTGTGTGGGAAGCAGTCAGTATAATTCCCGTCGTGTGGTGACGC  9300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9239  GTGAGGAGTCCTGGCGTGTGTGGGAAGCAGTCAGTATAATTCCCGTCGTGTGGTGACGC  9298

Qy  9301  CTCACGACGTATTGTCCGCTGTGCAGAGCGTAGTACCAAGGGCTGACCCCGTTTTTG  9360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9299  CTCACGACGTACTTGTCCGCTGTGCAGAGCGTAGTACCAAGGGCTGCGACCCCGGTTTTTG  9358

Qy  9361  TTCCAAGGGGAGGCAACCCCCGCTTGGAATTAAAAACT  9399
          ||||||||||||||||||||||||||||||||||||||
Db  9359  TTCCAAGCCGGAGGGCAACCCCCGCTTGGAATTAAAAACT  9397
```

RESULT 5

```
HGB277947
LOCUS       HGB277947               9397 bp    RNA     linear   VRL 01-MAR-2002
DEFINITION  Hepatitis GB virus B genomic RNA.
ACCESSION   AJ277947
VERSION     AJ277947.1  GI:13162187
KEYWORDS    ORF1; polyprotein.
SOURCE      Hepatitis GB virus B
  ORGANISM  Hepatitis GB virus B
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae.
REFERENCE   1
  AUTHORS   Sbardelli,A., Scarselli,E., Verschoor,E., De Tomassi,A.,
            Lazzaro,D. and Traboni,C.
  TITLE     Generation of infectious and transmissible virions from a GB virus
            B full-length consensus clone in tamarins
  JOURNAL   J. Gen. Virol. 82 (Pt 10), 2437-2448 (2001)
  MEDLINE   21446677
  PUBMED    11562537
REFERENCE   2  (bases 1 to 9397)
  AUTHORS   Traboni,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-2000) Traboni C., Biochemistry, IRBM P.
            Angeletti, via Pontina km. 30.600, 00040 Pomezia (Roma), ITALY
FEATURES             Location/Qualifiers
     source          1..9397
                     /organism="Hepatitis GB virus B"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:39113"
                     /clone="ACY/GBV-B/FL-3"
     5'UTR           1..445
     CDS             446..9040
                     /note="ORF1"
                     /codon_start=1
                     /product="polyprotein"
                     /protein_id="CAC33083.1"
                     /db_xref="GI:13162188"
                     /db_xref="GOA:Q999T0"
                     /db_xref="UniProt/TrEMBL:Q999T0"
                     /translation="MPVISTQFSPVPAPRTRKNKQTQASYPVSIKTSVERGQRAKRKV
                     QRDARPRNYKIAGIHDGLQTLAQAALPAHGWGRQDPRHKSRNLGILLDYPLGWIGDVT
                     THTPLVGPLVAGAVVRPVCQIVRLLEDGVNWATGWFGVHLFVVCLLSLACPCSGARVT
                     DPDTNTTILTNCCQRNQVIYCSPSTCLHEPGCVICADECWVPANPYISHPSNWTGTDS
                     FLADHIDFVMGALVTCDALDIGELCGACVLVGDWLVRHWLIHIDLNETGTCYLEVPTG
                     IDPGFLGFIGMAGKVEAVIFLTKLASQVPYAIATMFSSVHYLAVGALIYYASRGKWY
                     QLLALMLYIEATSGNPIRVPTGCSIAEFCSPLMIPCPCHSYLSENVSEVICSPKWT
                     RPVTLEYNNSISWYPYTIPGARGCMVKFKNNTWGGCCRIRNVPSYCTMGTDAVWNDTRN
                     DPVNSTLLPPERWARLPSTPPVVRGSWLQVPCGFYSDVKDLATGLIITKDKAWKNYQVL
                     YSATGALSLTGVTTKAVVLILGLCGSKYLILAYLCYLSLCFGRASGYPLRRPVLPSQS
                     YLQAGWDVLSKAQVAPFALIFFICCYLRCRLRYAALLGFVPMAAGLPLTFFVAAAAAQ
                     PDYDWWVRLLVAGLVLVAGRDRGHRIALLVGPMIVSPGHRIALLHLVTPASAGFDTEIIGG
                     LTIPPVVALVVMSRFGPFAHLLPRCALVNSYLWQRWENWPWNVTLRPERFFLVLVCFP
                     GATYDALVTFCVCHVALLCLTSSAASFFGTDSRVRAHRMLVRLGKCHAWYSHYVLKFF
                     LLVFGENGVYFYKHLHGDVLPNDPASKLPLQEPFFPEGKARVYRNEGRRLACGDTVD
                     GLPVVARLGDLVFAGLAMPPDGWAITAPFTLQCLSERGTLSAMAVVMTGIDPRTWTGT
                     IFRLGSLATSYMGFVCDNLVTRLGSLVEVNKSDDDYWCVCGALPMAVAKGSSGAPILCSSG
                     AGSLTRCSCGETKGYLVTRLGGLVEVNKSDDDYWCVCGALPMAVAKGSSGAPILCSSG
                     HVIGMFTAARNSGGSVSQIRVRPLVCAGYHPQYTAHATLDIKPTVPNEYSVQILIAPT
```

```
                                                                                     Qy   541 ATCTGTTGAAAGGGGACAACGAGCAAAGCGCAAAGTCCAGGCGCAAAGTCCTCGGCCTCGTAA  600
                                                                                              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                                                                                     Db   541 ATCTGTTGAAAGGGGACAACGAGCAAAGCGCAAAGTCCAGGCGCAAAGTCCTCGGCCTCGTAA  600

GSGKSTKLPLSYMQEKYEVLVLNPSVATTASMPKYMHATYGVNPNCYFNGKCNTGAS         Qy   601 TTACAAAATTGCTGTATCCATGATGGCTTGCAGACATTGGCTCAGCTGCTTTGCCAGC  660
LTYSTYGMLTGACSRNTYDVIICDECHATDATTVLGIGKVLTBAPSKNVRLVVLATAT                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
PPGVIPTPHANITBIQLTDEGTIPFHGKKIKEENLKKGRHLIFEATKKGHCDELANELA       Db   601 TTACAAAATTGCTGTATCCATGATGGCTTGCAGACATTGGCTCAGCTGCTTTGCCAGC  660
RKGITAVSYYRGCDISKLPEGDCVVVATDALCTGYTGDPDSVVDCSLMVEGTCHVDLD
PTFMGVRVCGVSAIVKGQRRGRTGRGRAGIYYYVDGSCTPSGMVPECNIVEAFDAAK         Qy   661 TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTCGCAATCTTGGAATCCTTCTGA  720
AWYGLSSTRAQTILDTYRTQPGLPAIGANLDEWADLFSMVNPEPSFVNTAKRTADNYV                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
LLTAAQLQLCHQYGYAAPNDAPRWQGARLGKKPCGVLWRLDGADACPGPEPSEVTRYQ       Db   661 TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTCGCAATCTTGGAATCCTTCTGA  720
MCFTEVNTSGTAALAVGVGVAMAYLAIDTFGATCVRRCWSITSVPTGATVAPVVDEBE
IVEECASPIFPLEAMVAAIDKLKSITTTTSPFTLETALEKLNTFLGPHAATILAIIEYC       Qy   721 TTACCCTTTGGGTGGTGATGTTACAACTCAGTTACAACTCAGTTGCTGCTGGAGGATGGAGTCAA  780
CGLVTLPDNPFASCVFAFIAGITTTPLPHKIKMPLSLFGGAIIASKLTDARGALAFMWAG               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AAGTALGTWTSVGFVFDMLGGYAAASSTACLTFKCLMGEWPTMDQLAGLIVYSAFNPAA       Db   721 TTACCCTTTGGGTGGTGATGTTACAACTCAGTTACAACTCAGTTGCTGCTGGAGGATGGAGTCAA  780
GVVGVLSACAMFALTTAGPDHWPNRLLTMLARSNTVCNEYFVICFNVLKAGVQSMVNIPGC
TPWSVISACIRWLHTPTEDDCGLIAWGLEIWQYVCNFFVICFNVLKAGVQSMVNIPGC        Qy   781 GGCAGGAGGCGGTCGTTCGACCAGTCTGCCAGATAGTACGCTTGCTGCTGGAGGATGACGGAGTCAA  840
PFYSCQKGYKGPNIGSGMLQARCPCGAELIFSVENGFAKLYKGPRTCSNYWRGAVPVN                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ARLCGSARPDPTDWTSLVVNYGVRDYCKYEKLGDHIFVTAVSSNVCFTQVPPTLRAA         Db   781 GGCAGGAGGCGGTCGTTCGACCAGTCTGCCAGATAGTACGCTTGCTGCTGGAGGATGACGGAGTCAA  840
VAVDGVQVQCYLGEPKTPWTTSACCYGPDGKGKTVKLPFRVDGHTPGVRMQLNLRDAL
ETNDCNSTNNTPSDEAAVSALIVFKQELRRTNQLLEAISAGVDTTKLPAPSIEEVVRK        Qy   841 CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTTGTGTATGTCGTCATCTTTGGCCTG  900
RQFRARTGSLTLPPPPRSVPGVSCPESLQRSDFLPSPSNLPPSPVLQLAMPMPLLGA                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GECNPFTAIGCAMTETGGPDDLPSYPPKKEVSEWSDESWSTATTASSYVTGPPYPKI        Db   841 CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTTGTGTATGTCGTCATCTTTGGCCTG  900
RGKDSTQSAPAKRPTKKKLGKSEFSCCSMSYTWTDVISFKTASKVLSAITRAITSGFLKQ
RSLVYVTEPRDAELRKQKVTINRQPLFPPSYHVGVRLAKEKASKVVGVMVWDVDEVAAH       Qy   901 TCCCTGTAGTGGGGCGGCGGGTCACTGACCCCAGACACAAAATACCACAAATCCTGACCAATTG  960
TPSKSAKSHITGLRGTDVRSGAARKAVLDLQKCVEAGEIPSHYRQTVIVPKEEVFVKT                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
PQKPTKKPPRLISYPHLEMRCVKMYYGQVAPDVVKAVMGDAYGFVDPRTRVKRLLSM        Db   901 TCCCTGTAGTGGGGCGGCGGGTCACTGACCCCAGACACAAATACCACAAATCCTGACCAATTG  960
WSPDAVGATCTVCFDSTITPEDIMVETDIYSAAKLSDQHRAGIHTIARQLYAGGPMI
AYDDGREIGYRRCRSSGVVTTSSSNSLTCWLKVNAAAEQAGMKNPRFLICGDDCTVIWK       Qy   961 CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCCTTCCGCCAATCCGTACAATCTTCACACCTTCCAA  1020
SAGADADKQAMRVFASWMKVMGAPODCVPQPKYSLEELTSCSSNVTSGITTKSGKPYYF               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
LTRDPRIPLGRCSAEGLGYNPSAAWIGYLIHHYPCLWVSRVLAVHFMEOMLFEDKLPE       Db   961 CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCCTTCCGCCAATCCGTACAATCTTCACACCTTCCAA  1020
TVTFDWYGKNYTVPVEDLPSIIAGVHGIEAFSVVRYTNAEILRVSQSLTDMTMPPLRA
WRKKARAVLASAKRRGGAHAKLARFLLWHATSRPLPDLDKTSVARYTTFNYCDVYSPE       Qy  1021 TGTGATCGTGCGCACGGACTCCTTCTTCGGTTCCCGCCAATCCGTACAATCTCACACCTTCCAA  1080
GDVFVTPQRRLQKFLVKYLAVIVFALGLIAVGLAIS"                                     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                                                                 Db  1021 TGTGATCGTGCGCACGGACTCCTTCTTCGGTTCCCGCCAATCCGTACAATCTCACACCTTCTGT  1080
3'UTR
                                                                 Qy  1081 TTGGACTGGCACGGACTCCTTCTTCGGTGTCGTTGTTGTTGTTGTTTATGGGCGCTCTTGT  1140
ORIGIN                                                                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
       9041.  .9397 .                                          Db  1081 TTGGACTGGCACGGACTCCTTCTTCGGTGTCGTTGTTGTTGTTGTTTATGGGCGCTCTTGT  1140

                                                                 Qy  1141 GACCTGTGACGCCCTTGACGTTGTGTGGTTGCGTTGTGTGTTATTAGTTGCGTGTGACTG  1200
Query Match   99.4%;  Score 9338.6;  DB 14;  Length 9397;              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Best Local Similarity  99.7%;  Pred. No. 0;                       Db  1141 GACCTGTGACGCCCTTGACGTTGTGTGGTTGCGTTGTGTGTTATTAGTTGCGTGTGACTG  1200
Matches 9368;  Conservative  0;  Mismatches 29;  Indels  2;  Gaps  1;
                                                                 Qy  1201 GCTTGTCAGGCACTGGCTTATTCATCATATGAACCTCAATGAAACTGGTACTTGTTACCTGA  1260
Qy    1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCTAG   60         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||       Db  1201 GCTTGTCAGGCACTGGCTTATTCATCATATGAACCTCAATGAAACTGGTACTTGTTACCTGA  1260
Db    1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCTAG   60
                                                                 Qy  1261 AGTGCCACTGGAATAGATCCTGGGTTTATCGGGTGGATGGCCCGGCAAGGT  1320
Qy   61 CAGGGGGCGTGGGGATTTCCCCGCCCGTCTGCAGAGAGGGTGGAGCCAACCACCTTAGTAT  120         |||||||||||||||||||||||||||||||||||||||||||||||||||
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||       Db  1261 AGTGCCACTGGAATAGATCCTGGGTTTATCGGGTGGATGGCCCGGCAAGGT  1320
Db   61 CAGGGGGCGTGGGGATTTCCCCGCCCGTCTGCAGAGAGGGTGGAGCCAACCACCTTAGTAT  120
                                                                 Qy  1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCAAGGACG  1380
Qy  121 GTAGGCGGCGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180          |||||||||||||||||||||||||||||||||||||||||||||||||
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||       Db  1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCAAGGACG  1380
Db  121 GTAGGCGGCGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180
                                                                 Qy  1381 GTTTAGCAGTGTACACTGCTCTGATCTACTACTATGTGT  1440
Qy  181 CCTGATGGGGTCATGGGTTCGGTGGTTGGTGGTGGTGGCCTTAGGCAGCCTCACGCGCCACCA  240         |||||||||||||||||||||||||||||||||||||||
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||   Db  1381 GTTTAGCAGTGTACACTGCTCTGATCTACTACTATGTGT  1440
Db  181 CCTGATGGGGTCATGGGTTCGGTGGTTGGTGGTGGTGGCCTTAGGCAGCCTCACGCGCCACCA  240
                                                                 Qy  1441 GTGGTATCAGTTGCTCTTGACCAAACTGCTTTACATAGAAGCGACCCCAT  1500
Qy  241 CCTCCCAGATAGAGCGCGGCACTGTAGGGAAGAAGAGACCGGGGACCGGTCACTACCAAGGACG  300       |||||||||||||||||||||||||||||||||||||||||||||||||
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||   Db  1441 GTGGTATCAGTTGCTCTTGACCAAACTGCTTTACATAGAAGCGACCCCAT  1500
Db  241 CCTCCCAGATAGAGCGCGGCACTGTAGGGAAGAAGAGACCGGGGACCGGTCACTACCAAGGACG  300
                                                                 Qy  1501 CAGGGTGCCCACTGAGTGCTCAATAGCTCAAAACAAGTCCTGTACC  1560
Qy  301 CAGAGCTCTTTTTGAGTATCACGCCCTCCGGAAGTAGTTGGCAGCCTCCACGCGCCACCA  360          ||||||||||||||||||||||||||||||||||||||||||||||
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||        Db  1501 CAGGGTGCCCACTGAGTGCTCAATAGCTCAAAACAAGTCCTGTACC  1560
Db  301 CAGAGCTCTTTTTGAGTATCACGCCCTCCGGAAGTAGTTGGCAGCCTCCACGCGCCACCA  360
                                                                 Qy  1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC  1620
Qy  361 TGGGATGGTTGGGGTTAGCCATCAATACCGTACCGCCTGTTAGGGGATGGTTCTTGCGAGGGGAT  420     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||       Db  1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC  1620
Db  361 TGGGATGGTTGGGGTTAGCCATCAATACCGTACCGCCTGTTAGGGGATGGTTCTTGCGAGGGGAT  420
                                                                 Qy  1621 CAGGCCTATCACTCCTAGAGTATAACAACTCCATATCTGGTACCCCTATACAATCCCTGG  1680
Qy  421 CTGGGAGTCTCGTAGACCGTAGCACATGCCCTGTATTTCTACTCAAACAAGTCCTGTACC  480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CTGGGAGTCTCGTAGACCGTAGCACATGCCCTGTATTTCTACTCAAACAAGTCCTGTACC  480

Qy  481 TGCGCCCAGAACGCGCAAGAGAACAAGCAGACGCAGGCTTCATTATCCGTGTCCATTAAAAC  540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TGCGCCCAGAACGCGCAAGAGAACAAGCAGACGCAGGCTTCATTATCCGTGTCCATTAAAAC  540
```

Right column (positions 4921–6060):

```
Qy  4921 TTATGTTTTGTTGACTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCC 4980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921 TTATGTTTTGTTGACTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCC 4980

Qy  4981 CAATGACGCACCACGGTGCAGGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG   5040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981 CAATGACGCACCACGGTGCAGGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG   5040

Qy  5041 GCGCTTGACGCGGCTGACGCGTCCTGGCCCAGAGCCCAGCGAGGTGACCAGATACCA     5100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041 GCGCTTGACGCGGCTGACGCGTCCTGGCCCAGAGCCCAGCGAGGTGACCAGATACCA     5100

Qy  5101 AATGTGCTTCACTGAGTCAATACTTCTGGGACAGCCGCACTCGCTGTTGGCCGTTGGGAGT 5160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101 AATGTGCTTCACTGAGTCAATACTTCTGGGACAGCCGCACTCGCTGTTGGCCGTTGGGAGT 5160

Qy  5161 GGCTATGGCTTATCTAGCCATTGACACTTTGGCGCCACTTGTGTGCGCGGTTGCTGGTC   5220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161 GGCTATGGCTTATCTAGCCATTGACACTTTGGCGCCACTTGTGTGCGCGGTTGCTGGTC   5220

Qy  5221 TATTACATCAGTCCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAAGAAATCGT  5280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221 TATTACATCAGTCCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAAGAAATCGT  5280

Qy  5281 GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGTTTGCTGCAATTGACAAGCTGAA  5340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281 GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGTTTGCTGCAATTGACAAGCTGAA  5340

Qy  5341 GAGTACAATCACCACAACTAGTCTTTCACATTGGAAACCGCCCTTAAACACTTAACAC    5400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341 GAGTACAATCACCACAACTAGTCTTTCACATTGGAAACCGCCCTTAAACACTTAACAC    5400

Qy  5401 CTTTCTTGGGCCTCATGGAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGTTTAGT   5460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401 CTTTCTTGGGCCTCATGGAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGTTTAGT   5460

Qy  5461 CACTTACCTGACAATCCCTTGCATCATGGTGTTTGCTTTCATTGGAGGCCGAATTGCGTCCAA 5520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461 CACTTACCTGACAATCCCTTGCATCATGGTGTTTGCTTTCATTGGAGGCCGAATTGCGTCCAA 5520

Qy  5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTTGGAGGGGCTGCGGGAACAGCTCT 5580
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTTGGAGGGGCTGCGGGAACAGCTCT 5580

Qy  5581 GCTTACAGACGCTAGAGGGCACTGGCGTTCATGATGCGGCTCGCGGGACATGCTGCCGCCTC 5640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581 GCTTACAGACGCTAGAGGGCACTGGCGTTCATGATGCGGCTCGCGGGACATGCTGCCGCCTC 5640

Qy  5641 TGGTACATGACATGCGTGGTTTGTCTTTGACATGCCTAGGCGGCTATCGCCGCCTC     5700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641 TGGTACATGACATGCGTGGTTTGTCTTTGACATGCCTAGGCGGCTATCGCCGCCTC     5700

Qy  5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGCCCACTATGGATCAGCT   5760
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGCCCACTATGGATCAGCT   5760

Qy  5761 TGCTGGTTAGTGCTACTCCGGGTTCAATCCGGCGGCCAGGAGTGTGGGCGTCTTGTCAGC  5820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761 TGCTGGTTAGTGCTACTCCGGGTTCAATCCGGCGGCCAGGAGTGTGGGCGTTTGTCAGC  5820

Qy  5821 TTGTGCAATGTTTGCTTTGACAACAGCAGGGCCAGATCACTGCCCAACAGACTTCTTAC  5880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821 TTGTGCAATGTTTGCTTTGACAACAGCAGGGCCAGATCACTGCCCAACAGACTTCTTAC  5880

Qy  5881 TATGCTTGCTGCTAGGAGCAACACTGTATGTAATGACACTGTATTGCCACTCGTGACATCCG 5940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881 TATGCTTGCTGCTAGGAGCAACACTGTATGTAATGACACTGTATTGCCACTCGTGACATCCG 5940

Qy  5941 CAGGAAGATACTGGGCATTCTGGGGCATCTCACCCCCTGGAGTGTCATATCAGCTTGCAT   6000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941 CAGGAAGATACTGGGCATTCTGGGGCATCTCACCCCCTGGAGTGTCATATCAGCTTGCAT   6000

Qy  6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGCGCTCATTGCTTCGGGGTCTAGAGAT 6060
```

Left column (positions 3841–4920):

```
Qy  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGCAAC     3900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGCAAC     3900

Qy  3901 CAAATTACCACTTCTTACATGCAGGAGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT   3960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 CAAATTACCACTTCTTACATGCAGGAGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT   3960

Qy  3961 GGCTACAACAGCATCAATGCCAAAGTACATGCACGGCGTACGGGCTGAATCCAAATTG    4020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 GGCTACAACAGCATCAATGCCAAAGTACATGCACGGCGTACGGGCTGAATCCAAATTG    4020

Qy  4021 CTATTTTAATGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGCAT   4080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 CTATTTTAATGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGCAT   4080

Qy  4081 GTACCTGACCGGAGCATGTTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGC  4140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 GTACCTGACCGGAGCATGTTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGC  4140

Qy  4141 TACCGATGCAACCACCGTGTGGGCATTGGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA   4200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 TACCGATGCAACCACCGTGTGGGCATTGGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA   4200

Qy  4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCTGGAGTAATCCCTGGAGTAATGCCACACAGC 4260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCTGGAGTAATCCCTGGAGTAATGCCACACAGC 4260

Qy  4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAGAT   4320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAGAT   4320

Qy  4321 TAAGGAGGAAAATCTGAAGAAGGGAGACACACTTATCTTTGAGGCTACCAAAAAAACACTG  4380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321 TAAGGAGGAAAATCTGAAGAAGGGAGACACACTTATCTTTGAGGCTACCAAAAAAACACTG  4380

Qy  4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG  4440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG  4440

Qy  4441 ATGTGACATCTCAAAAATCCCTGAGGCGCGACTGTGTAGTAGTGCCACTGAGTGCGTTGTG 4500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441 ATGTGACATCTCAAAAATCCCTGAGGCGCGACTGTGTAGTAGTGCCACTGAGTGCGTTGTG 4500

Qy  4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTGCAGCCTCATGTGTAGAGGCAC  4560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTGCAGCCTCATGTGTAGAGGCAC  4560

Qy  4561 ATGCCATGTGACCTTGACCCTACTTTCACCATGGTGTCGTGTGTCCACTGAGTTCGCGGGTTTCAGC 4620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561 ATGCCATGTGACCTTGACCCTACTTTCACCATGGTGTCGTGTGTCCACTGAGTTCGCGGGTTTCAGC 4620

Qy  4621 AATAGTTAAAGGCCAGCGTAGGGGCCGTGGGAGAGCGGTGGCATATACTACTA      4680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621 AATAGTTAAAGGCCAGCGTAGGGGCCGTGGGAGAGCGGTGGCATATACTACTA      4680

Qy  4681 TGTGTGGTTTAGCTACTCCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT      4740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681 TGTGTGGTTTAGCTACTCCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT      4740

Qy  4741 CGACGCAGCCAAGGCCATGGCATGGTATGGTTTTGTCATCAACAGAAGCTATTCTGACAC  4800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741 CGACGCAGCCAAGGCCATGGCATGGTATGGTTTTGTCATCAACAGAAGCTATTCTGACAC  4800

Qy  4801 CTATCGCACCCACCCCAACCTGGGTTACCTGGGGATAGGAGCAAATTTGGACGAGTGGGCTGATCT 4860
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801 CTATCGCACCCACCCCAACCTGGGTTACCTGGGGATAGGAGCAAATTTGGACGAGTGGGCTGATCT 4860

Qy  4861 CTTTTCTAATGGTCAACCCGAACCTTCATTTGTCAATACTGCAAATACTGCTGACAA    4920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861 CTTTTCAATGGTCAACCCGAACCTTCATTTGTCAATACTGCAAATACTGCTGACAA    4920
```

The body of this page consists of a two-block DNA sequence alignment (Query "Qy" vs Database "Db"), with position numbers running from 6001 to 8220. Each aligned pair is shown with a row of vertical match bars ("|") between the two sequences.

Left column block (positions 6001–7140):

```
Db  6001  ||||||||||||||||||CCGTTGGCTCCACACCCCGACGGAGGAGTGATTGCGGCCTCATTGGGCTTAAGTGCTTGGGGTCTAGAGAT  6060
Qy  6061  TTGGCAGTATGTGTGCAATTCTTCTTGTGATTTCGCTTTAATGTCCTTAAAGCTGGAGTTCA  6120
Db  6061  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TTTGGCAGTACGTGTGCAATTTCTTGTGATTTCGCTTTAATGTCCTTAAAGCTGGAGTTCA  6120
Qy  6121  GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC  6180
Db  6121  GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC  6180
Qy  6181  CTGGGATTGGATCAGGTATGCTCCAAGCACGCTGTCCATGCGGGTGCTGAACTCATCTTTTC  6240
Db  6181  CTGGGATTGGATCAGGTATGCTCCAAGCACGCTGTCCATGCGGGTGCTGAACTCATCTTTTC  6240
Qy  6241  TGTTGAGAATGGTTTTGCAAAACTTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300
Db  6241  TGTTGAGAATGGTTTTGCAAAACTTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300
Qy  6301  AGGGGCTGTTCCAGTCAACGCTAGGCTGTGTGTGGCTGTGTGGGCTGGCTAGGACCGACCCAACTGATTG  6360
Db  6301  AGGGGCTGTTCCAGTCAACGCTAGGCTGTGTGTGGCTGTGTGGGCTGGCTAGGACCGACCCAACTGATTG  6360
Qy  6361  GACTAGTCTGTCGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA  6420
Db  6361  GACTAGTCTGTCGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATTGGGAGA  6420
Qy  6421  TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCAAC  6480
Db  6421  TCACATCTTTGTTGTCGCGGTTACACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCAAC  6480
Qy  6481  CTTGAGAGCTGCAGTGGCCGTGACGGCGTACAGGTTCAGGTTTATCTGAGCCGCAGT  6540
Db  6481  CTTGAGAGCTGCAGTGGCCGTGACGGCGTACAGGTTCAGGTTTATCTGAGCCGCAGT  6540
Qy  6541  AACTCCTTGGACACAGACATCTGCTTGCTGTCTGGTCCTGACGCGGTAAGGGTAAAACTGTTAA  6600
Db  6541  AACTCCTTGGACACAGACATCTGCTTGCTGTCTGGTCCTGACGCGGTAAGGGTAAAACTGTTAA  6600
Qy  6601  GCTTCCCTTCCGGGTTGACGGTCACACACCTGGTGTGCGCATGCAACTTAATTTGCGTGA  6660
Db  6601  GCTTCCCTTCCGGGTTGACGGTCACACACCTGGTGTGCGCATGCAACTTAATTTGCGTGA  6660
Qy  6661  TGCACTTGAGACAAATGACTGTAATTCACAAAACAACACTCCTAGTGATGAGCCGCAGT  6720
Db  6661  TGCACTTGAGACAAATGACTGTAATTCACAAAACAACACTCCTAGTGATGAGCCGCAGT  6720
Qy  6721  GTCCGGCTCTTCGTGTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC  6780
Db  6721  GTCCGGCTCTTCGTGTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC  6780
Qy  6781  AGCTGGGCGTTGACACCACCTGTTCTCACAGTTGGCCAGTGCCCATGCGAAGAGGTAGTTGGTAAGAAA  6840
Db  6781  AGCTGGGCGTTGACACCACCTGTTCTCACAGTTGGCCAGTGCCCATGCGAAGAGGTAGTTGGTAAGAAA  6840
Qy  6841  GCGCCAGTTCCGGGCAAGAACTGTTCGCTTACCTGCGTTCGCTTACCTGCGAGATCCGTCCC  6900
Db  6841  GCGCCAGTTCCGGGCAAGAACTGTTCGCTTACCTGCGTTCGCTTACCTGCGAGATCCGTCCC  6900
Qy  6901  AGGAGTGTGCATGTCCTGAAAGCCTGCAACGAAGCCTGCGAACGAAGTGACCCGGTTAGAAGGTTCCTTCAAACCT  6960
Db  6901  AGGAGTGTGCATGTCCTGAAAGCCTGCAACGAAGCCTGCGAACGAAGTGACCCGGTTAGAAGGTTCCTTCAAACCT  6960
Qy  6961  CCCTTCCTTCACCACCTGTTCTCACAGTTGGCCATGCCGATGCCCGTTGGGAGCGGGTGA  7020
Db  6961  CCCTTCCTTCACCACCTGTTCTCACAGTTGGCCATGCCGATGCCCGTTGGGAGCGGGTGA  7020
Qy  7021  GTGTAACCCTTTCACTGAGTGTCAATGACCCGAAACAGGCGGCCCTGATGA  7080
Db  7021  GTGTAACCCTTTCACTGAGTGTCAATGACCCGAAACAGGCGGCCCTGATGA  7080
Qy  7081  TTTACCAGTTACCCTCCCAAAAAGGAGGTCTCGAATGGTCAGACGAAAGTTCGCAC  7140
Db  7081  ||||||||||||||||||
```

Right column block (positions 7081–8220):

```
Db  7081  TTTACCAGTTACCCTCCCAAAAAGGAGGTCTCGAATGGTCAGACGAAAGTTCGTCAAC  7140
Qy  7141  GGCTACAACCGCTTCCAGCTACGTACTGGCCCCCGTACCCTAAGATACGGGGAAAGGA  7200
Db  7141  GGCTACAACCGCTTCCAGCTACGTACTGGCCCCCGTACCCTAAGATACGGGGAAAGGA  7200
Qy  7201  TTCCACTCAGTCAGCCCCGCCAAAGCCTACAAAAAGAAGTTGGGAAAGAGTGAGTT  7260
Db  7201  TTCCACTCAGTCAGCCCCGCCAAAGCCTACAAAAAGAAGTTGGGAAAGAGTGAGTT  7260
Qy  7261  TTCGTGAGCATGAGCTGGACCGACGTGATTAGTAGCTTCAAAACTGCTTCTAAAGT  7320
Db  7261  TTCGTGAGCATGAGCTGGACCGACGTGATTAGTAGCTTCAAAACTGCTTCTAAAGT  7320
Qy  7321  TCTGTCTGCAACTCGGGCCATCACTAGTGTGTTCCTCAAACAAGATCATTGGTGTATGT  7380
Db  7321  TCTGTCTGCAACTCGGGCCATCACTAGTGTGTTCCTCAAACAAGATCATTGGTGTATGT  7380
Qy  7381  GACTGAGCCGCGGGATGCGGGAGCTTAGAAACAAAAAGTCACTATTAATAGACAACCTCT  7440
Db  7381  GACTGAGCCGCGGGATGCGGGAGCTTAGAAACAAAAAGTCACTATTAATAGACAACCTCT  7440
Qy  7441  GTTCCCCCATCATACACAAGCAAGCAAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT  7500
Db  7441  GTTCCCCCATCATACACAAGCAAGCAAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT  7500
Qy  7501  CGGTGTCATGTGGACTATGGACAGCTCCACACGCCCTCTAAGTCTCGTAAGTC  7560
Db  7501  CGGTGTCATGTGGACTATGGACAGCTCCACACGCCCTCTAAGTCTCGTAAGTC  7560
Qy  7561  CCACATCACTGGCCTCGGGGGCACTGATGTTCGTTCGTTCTGGAGCAGCCGCAAGGCTGTTCT  7620
Db  7561  CCACATCACTGGCCTCGGGGGCACTGATGTTCGTTCGTTCTGGAGCAGCCGCAAGGCTGTTCT  7620
Qy  7621  GGACTTGCAGAAGTGTGTCAGGCAGGTGAGTACCGAGTCATTATCGGCAAACTGTGAT  7680
Db  7621  GGACTTGCAGAAGTGTGTCAGGCAGGTGAGTACCGAGTCATTATCGGCAAACAGTGAT  7680
Qy  7681  AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAACAACAAAGAAAACCCCCAAG  7740
Db  7681  AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAACAACAAAGAAAACCCCCAAG  7740
Qy  7741  GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAAGTGTACTACGGTCAGGT  7800
Db  7741  GCTCATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAAGTGTACTACGGTCAGGT  7800
Qy  7801  TGCTCCTGACGTAGTTGAAAGCTGTCATGGGAGTGGTTGTAGATCCGGGTTGTAGATCCACGTAC  7860
Db  7801  TGCTCCTGACGTAGTTGAAAGCTGTCATGGGAGTGGTTGTAGATCCGGGTTGTAGATCCACGTAC  7860
Qy  7861  CCGTGTCAAGCGTCTGTTGTCGATGTGTGCGAGTCGGATGCGAGCCGACATGCGGATAC  7920
Db  7861  CCGTGTCAAGCGTCTGTTGTCGATGTGTGCACCCGAGTCGGATGCGAGCCACCATGCGGATAC  7920
Qy  7921  AGTGTGTTTGACAGTACCATCACACCCGAGATATCATGGTGAAGCAGACATCTACTC  7980
Db  7921  AGTGTGTTTGACAGTACCATCACACCCGAGATATCATGGTGAAGCAGACATCTACTC  7980
Qy  7981  AGCAGCTAAACTCAGTGACCTGCCAACACCGAGCTGGCCCAGCAGGCAGTTATA  8040
Db  7981  AGCAGCTAAACTCAGTGACCTGCCAACACCGAGCTGGCCCAGCAGGCAGTTATA  8040
Qy  8041  CGCTGGGAGGACCGATGATCGTTATGATGGCCGAGAGATCGGGATATCGTAGGTGTAGGTC  8100
Db  8041  CGCTGGGAGGACCGATGATCGTTATGATGGCCGAGAGATCGGGATATCGTAGGTGTAGGTC  8100
Qy  8101  TTCCGGGTCTATACTACTTCAAGTTCAACAGTTTGACCTGCTCGGCTGAAGGTAAATGC  8160
Db  8101  TTCCGGGTCTATACTACTTCAAGTTCAACAGTTTGACCTGCTCGGCTGAAGGTAAATGC  8160
Qy  8161  TGCAGCCGAACAGGCTGGCATGGCATGATGAGGAACCCTCGCTTCCTTATTTGCGGCATGATTGCAC  8220
Db  8161  TGCAGCCGAACAGGCTGGCATGGCATGATGAGGAACCCTCGCTTCCTTATTTGCGGCATGATTGCAC  8220
```

```
                                                              Qy   9301 CTCACGACGTATTTGTCCGCTGTGCAGAGCGTAGTACCAAGGGCTGACCCGGTTTTTG 9360
                                                                        ||||||||      ||||||||||||||||||||||||||||||||||||||||||||||
                                                              Db   9299 CTCACGACGTACTTGTCCGCTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTG 9358

                                                              Qy   9361 TTCCAAGCGGAGGGCAACCCCCGCTTGGAATTAAAAACT 9399
                                                                        ||||||||||||||||||||||||||||||||
                                                              Db   9359 TTCCAAGCGGAGGGCAACCCCCGCTTGGAATTAAAAACT 9397
```

RESULT 6
AX057394
LOCUS   AX057394   9139 bp   DNA   linear   PAT 17-JAN-2001
DEFINITION   Sequence 3 from Patent WO0075337.
ACCESSION   AX057394
VERSION   AX057394.1 GI:12310133
KEYWORDS   .
SOURCE   Hepatitis GB virus B
  ORGANISM   Hepatitis GB virus B
                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae.
REFERENCE   1
  AUTHORS   Bukh,J., Yanagi,M., Emerson,S.U. and Purcell,R.H.
  TITLE   Infectious cdna clone of gb virus b and uses thereof
  JOURNAL   Patent: WO 0075337-A 3 14-DEC-2000;
                  THE GOVERNMENT OF THE UNITED STATES OF AMERICA (US)
FEATURES   Location/Qualifiers
     source   1..9139
                  /organism="Hepatitis GB virus B"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:39113"
ORIGIN

Query Match   97.2%;   Score 9133; DB 6; Length 9139;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9133;   Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   1 ACCACAAACACTCAGTTGTTACACTCCGTAGGAATGCTCTCCGGAGCACCCCCCTAG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCACAAACACTCAGTTTGTTACACTCCGTAGGAATGCTCTCCGGAGCACCCCCCTAG 60

Qy  61 CAGGGGCGTGGGGGATTCCCCGTCGCCCCGTCGCGCTCGCAGAAGGGTGAGCCAACCACCTAGTAT 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 CAGGGGCGTGGGGGATTTCCCCGTCGCCCCGTCGCGCTCGCAGAAGGGTGAGCCAACCACCTAGTAT 120

Qy 121 GTAGGCGGCGGGACTCATGACGCTCGCGCGTGGTGGCGGCGCAAGCGCGCAAGCGTTGGATGGC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GTAGGCGGCGGGACTCATGACGCTCGCGCGTGGTGGCGGCGCAAGCGCGCAAGCGTTGGATGGC 180

Qy 181 CCTGATGGCGTTCATGGGTTCATGGGTTCGGTGGTCGCGCTTAGGCAGCCTCCACGCCACCA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 CCTGATGGCGTTCATGGGTTCATGGGTTCGGTGGTCGCGCTTAGGCAGCCTCCACGCCACCA 240

Qy 241 CCTCCCAGATAGAGCGGCGGCACTGTAGGGGAAGACCGGGACCGGTCACTACCAAGGACG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 CCTCCCAGATAGAGCGGCGGCACTGTAGGGGAAGACCGGGACCGGTCACTACCAAGGACG 300

Qy 301 CAGACCTCTTTTTGAGTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTGT 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 CAGACCTCTTTTTGAGTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTGT 360

Qy 361 TGGGATGTTGGGTTAGCCTGACCGTACGTGCCTGATAGGGTCCTTGCGAGGGGAT 420
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 TGGGATGTTGGGTTAGCCTGACCGTACGTGCCTGATAGGGTCCTTGCGAGGGGAT 420

Qy 421 CTGGGAGTCTCTAGACCGTAGCACATGGCCTGTTATTCTACTCAAACAAGTCCTGTACC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 CTGGGAGTCTCTAGACCGTAGCACATGGCCTGTTATTCTACTCAAACAAGTCCTGTACC 480

Qy 481 TGCGCCCAGAACGCGCAAGAACAAGACGAGACGCAGCTTCATATCCTGTGTCATTAAAAC 540
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 TGCGCCCAGAACGCGCAAGAACAAGACGAGACGCAGCTTCATATCCTGTGTCATTAAAAC 540

Qy 541 ATCTGTTGAAAGGGGACACAGGGACACAAGCGCAAAGTCCAGCGACGATGCGCGTCGGCCTCGTAA 600
```

```
Qy  8221 CGTAATTTGGAAGAGCCCCGAGCGAGCGATGCGACAAACAAGCAATGCGTGTCTTTGCTAG 8280
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8221 CGTAATTTGGAAAAGCGCCGAGCGAGCGATGCGACAAACAAGCAATGCGTGTCTTTGCTAG 8280

Qy  8281 CTGGATGAAGGTGATGGTGACCACAGATTGTGTCCTCAACCCAAATACAGTTTGGA 8340
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281 CTGGATGAAGGTGATGGTGACCACAGATTGTGCCTCAACCCAAATACAGTTTGGA 8340

Qy  8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400

Qy  8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTGCCAGGGGTCTGGG 8460
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTGCCAGGGGTCTGGG 8460

Qy  8461 ATACAACCCCAGTGCTGCGTGGGTATCTAATACATCATCACTACCCCATGTTTGTGGGT 8520
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8461 ATACAACCCCAGTGCTGCGTGGGTATCTAATACATCATCACTACCCCATGTTTGTGGGT 8520

Qy  8521 TAGCCGTGTGTTGGCTGTCCATTTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8521 TAGCCGTGTGTTGGCTGTCCATTTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580

Qy  8581 GACTGTGACCTTTGACTGTGACCGGTATGGGAATAAATTATACGGTGCCTGTAGAAGAATCTGCCCAG 8640
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db  8581 GACTGTGACCTTTGACTGTGACCGGTATGGGAATAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640

Qy  8641 CATCATTGCTCGTGTCGTGTGCACGGTATTGAGGGCTTCTTCTCCGGTGGTGGTGCGCTACACCAACGCTGA 8700
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8641 CATCATTGCTCGTGTCGTGTGCACGGTATTGAGGGCTTCTTCTCCGGTGGTGGTGCGCTACACCAACGCTGA 8700

Qy  8701 GATCCTTCAGAGTTTCCCAATCACTACAACAGACATGACGACCATGCCCCGCCCCGCGAGCCTGGCG 8760
           |||||||| |||||||||||||||||||||||||||||||||||||||
Db  8701 GATCCTTCAGAGTTTCCCAATCACTACAACAGACATGACGACCATGCCCCGCCCCGCGAGCCTGGCG 8760

Qy  8761 AAAGAAAGCCAGGGCGTCCTCGCCCAGCGCAAGAGGCGTGGCGGAGCGTGGCGGAGCACACGGCAAAATT 8820
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8761 AAAGAAAGCCAGGGCGTCCTCGCCCAGCGCAAGAGGCGTGGCGGAGCGTGGCGGAGCACACGGCAAAATT 8820

Qy  8821 GGCTCGCTTCCTTCTTCTCTCGGCATGCTACCTCTAGACCTCTACCGACCTCTACCAGATTTGGATAAGACGAG 8880
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db  8821 GGCTCGCTTCCTTCTTCTCTCGGCATGCTACCTCTAGACCTCTACCGACCTCTACCAGATTTGGATAAGACGAG 8880

Qy  8881 CGTGGCTCGGTACACCACCTTTCAATATTGTGATGTTACTCCCCGAGGGGATGTGTT 8940
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db  8881 CGTGGCTCGGTACACCACCTTTCAATATTGTGATGTTACTCCCCGAGGGGATGTGTT 8940

Qy  8941 TATTACACCAGAGAAGATTGCAGAAGTTCAGAAGTTTCTTGTGAAGTATTGGCTGTCATTGTTTT 9000
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db  8941 TGTTACACCAGAGAAGATTGCAGAAGTTCAGAAGTTTCTTGTGAAGTATTGGCTGTCATTGTTTT 9000

Qy  9001 TGCCCTAGGGCTCATTGCTCGTGTTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA 9060
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db  9001 TGCCCTAGGGCTCATTGCTCGTGTTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA 9060

Qy  9061 CTAACAGTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCCGCAACAGGGGGACCCCGGGC 9120
           ||||||   ||||||||||||||||||||||||||||||||||||||
Db  9061 CTAACAG--TTTTTTTTTTTTTTTTTTTAGGGCAGCCGCAACAGGGGGACCCCGGGC 9118

Qy  9121 TTAACGACCCCGCGATGTGAGTTTGGCGACCATGGTGGATCGAGATCGAGAACCGGTTCGGGTGAA 9180
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db  9119 TTAACGACCCCGCGATGTGAGTTTGGCGACCATGGTGGATCGAGATCGAGAACCGGTTCGGGTGAA 9178

Qy  9181 GCCATGTCTCGAAGGGGATGACGTCCCTTCGCTCCATCCACAAAAACCGTCTCGGGTGTG 9240
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db  9181 GCCATGTCTCGAAGGGGATGACGTCCCTTCGCTCCATCCACAAAAACCGTCTCGGGTGTG 9238

Qy  9241 GTGAGGAGTCCTGGCTGTGTGTGGGGAAGCAGTCAGTATAATTCCCGTCGTGGTGGTGACGC 9300
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db  9241 GTGAGGAGTCCTGGCTGTGTGTGGGGAAGCAGTCAGTATAATTCCCGTCGTGGTGGTGACGC 9298
```

```
Db   541  ATCTGTTGAAGGGGACAACGAGCAAAGCGCAAAGTCCAGCGCGATGCTCGGCGCCTCGTAA  600

Qy   601  TTACAAAATTGCTGTGGTATCCGATGGCTTGCGAGACATTGGCTCGAGACTCTTGCCAGC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  TTACAAAATTGCTGTGGTATCCGATGGCTTGCGAGACATTGGCTCGAGACTCTTGCCAGC  660

Qy   661  TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTTCGCAATCTTCTGAA         720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTTCGCAATCTTCTGA          720

Qy   721  TTACCCTTGGGGTGATTGGTGATGTTACAACTCACACACTCCTAGTAGGCCCGCTGGT    780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  TTACCCTTGGGGTGATTGGTGATGTTACAACTCACACACTCCTAGTAGGCCCGCTGT     780

Qy   781  GGCAGGAGCGGTCGTTGCGACCAGTCTGCCAGATAGTACGCTTGCTGGGGGATGGAGTCAA   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  GGCAGGAGCGGTCGTTGCGACCAGTCTGCCAGATAGTACGCTTGCTGGGGGATGGAGTCAA   840

Qy   841  CTGGGCTACTCGGTTGGTTCGGTGTCGGGTCGCTATCTTTGGCCTG                900
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   841  CTGGGCTACTCGGTTGGTTCGGTGTCGGGTCGCTATCTTTGGCCTG                900

Qy   901  TCCCTGTAGTGGGCGCGGGGTCACTGACCACCGGACCAACATACCACCAATTCCTGACCAATTG   960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  TCCCTGTAGTGGGCGCGGGGTCACTGACCACCGGACCAACATACCACCAATTCCTGACCAATTG   960

Qy   961  CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCTTCCACTTGCTCCACGAGCCTGGTTG    1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCTTCCACTTGCTCCACGAGCCTGGTTG    1020

Qy  1021  TGTGATCTGTGCGACGAGTGCTGGGTTCCCGCCAATCCGTACATCCACCCCTTCCAA     1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  TGTGATCTGTGCGACGAGTGCTGGGTTCCCGCCAATCCGTACATCCACCCCTTCCAA     1080

Qy  1081  TTTGGACTGGCACGGACACTCCTTCTGGCTGACCACATTGATTTTGTATGGGGCCTCTTGT   1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  TTTGGACTGGCACGGACACTCCTTCTGGCTGACCACATTGATTTTGTATGGGGCCTCTTGT   1140

Qy  1141  GACCTGTGACGCCCTTGACATTGGTGAGTTGTGTCGTGTGTATTAGTGGTATTAGTGGTGACTG   1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  GACCTGTGACGCCCTTGACATTGGTGAGTTGTGTCGTGTGTATTAGTGGTATTAGTGGTGACTG   1200

Qy  1201  GCTTGTCAGGCACTGGCTTATTCACATGAAACTGGTACTTGTTACCTGGA            1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GCTTGTCAGGCACTGGCTTATTCACATGAAACTGGTACTTGTTACCTGGA            1260

Qy  1261  AGTGCCCACTGGAATAGATCCTGGGTTTATCGGGTGGATGGCCGGCAAGGT           1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  AGTGCCCACTGGAATAGATCCTGGGTTTATCGGGTGGATGGCCGGCAAGGT           1320

Qy  1321  CGAGGCGTACGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATTGCGACTAT      1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  CGAGGCGTACGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATTGCGACTAT      1380

Qy  1381  GTTTAGCAGTGTACACTACCTGGTTGGCGCTCTGATCTACTATGCCTCTCGGGGGCAA    1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  GTTTAGCAGTGTACACTACCTGGTTGGCGCTCTGATCTACTATGCCTCTCGGGGGCAA    1440

Qy  1441  GTGTTGGTATCAGTTGCTCTCCTAGCTTTATCCATGAGGAGCGACACCTCTGGAAACCCCAT   1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  GTGGTGGTATCAGTTGCTCTCCTAGCTTTTACATAGAAGCGACACCTCTGGAAACCCCAT   1500

Qy  1501  CAGGGTGCCACTGGATCGTCAATAGCTGAGTTTGCTCGCCTTTGATGATACCCATGTCC    1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  CAGGGTGCCCACTGGATCGTCAATAGCTCAGTTTGCTCGCCTTTGATGATACCCATGTCC    1560

Qy  1561  TTGCCACTCTTATTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC     1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  TTGCCACTCTTATTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC     1620

Qy  1621  CAGGGCCTATCACTTCTAGAGTATAAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG   1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  CAGGGCCTATCACTTCTAGAGTATAAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG   1680
```

```
Qy  1681  TGGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCGTATTCGCAA   1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  TGGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCGTATTCGCAA   1740

Qy  1741  TGTGCCATCGTACTGCACTATGGCACTGATGCAGTGTGGAACGACACTCGCAACACTTA    1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  TGTGCCATCGTACTGCACTATGGCACTGATGCAGTGTGGAACGACACTCGCAACACTTA    1800

Qy  1801  CGAAGCATGCGGTGTAACACCATGGCTAACACCGCTAACAACCGCTCAGCCCTGAA      1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  CGAAGCATGCGGTGTAACACCATGGCTAACACCGCTAACAACCGCTCAGCCCTGAA      1860

Qy  1861  ATTGGCTATATTACAATACCCTGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC     1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  ATTGGCTATATTACAATACCCTGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC     1920

Qy  1921  AGGCCATTGTGTATTTGTGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAA   1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  AGGCCATTGTATTTTTTGAGGGATCAGATCAGATCAGTTTACTTTTATGACCCTGTGAA   1980

Qy  1981  TTCCACTCTCCTACCACCCGGAGAGTGGCTAGGTTGCCCGGTACCCCACCTGTGGTACG    2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  TTCCACTCTCCTACCACCCGGAGAGTGGCTAGGTTGCCCGGTACCCCACCTGTGGTACG    2040

Qy  2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACTAGGCACAGG     2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACTAGGCACAGG     2100

Qy  2101  ATTGATCACCAAAGACAAAGCCTGGAAAAAATTATCAGGTCTTATATTCCGCCACGGGATGC   2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  ATTGATCACCAAAGACAAAGCCTGGAAAAAATTATCAGGTCTTATATTCCGCCACGGGATGC   2160

Qy  2161  TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTTGGTGCTCAATTCGTTGTTGGGGTGTGTGG   2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTTGGTGCTCAATTCGTTGTTGGGGTGTGTGG   2220

Qy  2221  CAGCAAGTATCTTATTTTAGCCTACCTCTGTTGTTACTTGTCCCTTGTGTTTGGGCGCGCTTC   2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  CAGCAAGTATCTTATTTTAGCGTACCTCTGTTGTTACTTGTCCCTTGTGTTTGGGCGCGCTTC   2280

Qy  2281  TGGGTTACCCTTGCGTCCTGTGCGTCCATCGTCGTATCTCGTATCGTTGGCTGGGGATGT    2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TGGGTTACCCTTGCGTCCTGTGCGTCCATCGTCGTATCTCGTATCGTTGGCTGGGGATGT    2340

Qy  2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTTGATTTTCATCGTGTTGCTATCTCCG       2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  TTTGTCTCAAGCTCAAGTAGCTCCTTTTGATTTTCATCGTGTTGCTATCTCCG        2400

Qy  2401  CTGCAGGCTACGTTATGTCGAGCAGCTGCTGCCATGGCGTGTGCGGGCTTGCCCCT      2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  CTGCAGGCTACGTTATGTCGAGCAGCTGCTGCCATGGCGTGTGCGGGCCTTGCCCCT      2460

Qy  2461  AACTTTCTTTGTTGTTGCAGCAGCTGCTGCCAACCAGAATTATGACTGGTGGTGCGACTGCT    2520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  AACTTTCTTTGTTGTTGCAGCAGCTGCTGCCAACCAGAATTATGACTGGTGGTGCGACTGCT    2520

Qy  2521  AGTGGCAGGGTTAGTTTGTGGGCCGGCCGTAACCGTGGTCGTACATCTCGTCCTTGT     2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  AGTGGCAGGGTTAGTTTGTGGGCCGGCCGTAACCGTGGTCGTACATCTCGTCCTTGT     2580

Qy  2581  AGGTCCTTGGCCTTCGGTACGCGTTTGGTCATTTGGTTACCGCTGCTTCAGC         2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  AGGTCCTTGGCCTTCGGTACGCGTTTGGTCATTTGGTTACCGCTGCTTCAGC         2640

Qy  2641  TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGTTGTCAT        2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGTTGTCAT        2700

Qy  2701  GTCTCGTTTTGGCTTCTTTGTTCACTCGTGTTTAGTTAACTCCTATCT         2760
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  GTCTCGTTTTGGCTTCTTTGTTCACTCGTGTTTAGTTAACTCCTATCT         2760
```

```
Qy  2761 TTGGCAACGTTGGGAGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTTTCCT 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 TTGGCAACGTTGGGAGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTTTCCT 2820

Qy  2821 TGTGCTGTTGTTTGTTTCCCCGGTGCGACATATGACGGCGCTGGTGACTTTCTGTGTGTCA 2880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821 TGTGCTGTTGTTTGTTTCCCCGGTGCGACATATGACGGCGCTGGTGACTTTCTGTGTGTCA 2880

Qy  2881 CGTAGCTCTTCTATGTTTAAACATCCAGTGCAGCATCCGTTCGTTCTTGGGACTCTCAGGGT 2940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881 CGTAGCTCTTCTATGTTTAAACATCCAGTGCAGCATCCGTTCGTTCTTGGGACTCTCAGGGT 2940

Qy  2941 TAGGGCCATAGAAGTTGGTGCGTCTCGGAAAGTGTCATGCTTGGTGATTCTCATTATGT 3000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 TAGGGCCATAGAAGTTGGTGCGTCTCGGAAAGTGTCATGCTTGGTGATTCTCATTATGT 3000

Qy  3001 TCTTAAGTTTTTCCTCTTAGTGTTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 TCTTAAGTTTTTCCTCTTAGTGTTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA 3060

Qy  3061 TGGTGATGTCTTGCCTAATGATTTTGCTGAAACTACCATTTTCTCGAAGAGCCATTTTCCC 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 TGGTGATGTCTTGCCTAATGATTTTGCTGAAACTACCATTTTCTCGAAGAGCCATTTTCCC 3120

Qy  3121 TTTTGAAGGCAAGGCAAGGGGTCTATAGGAATGAAGAAGACGCTTGGCGTGTGGGGACAC 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 TTTTGAAGGCAAGGCAAGGGGTCTATAGGAATGAAGAAGACGCTTGGCGTGTGGGGACAC 3180

Qy  3181 GGTTGATGGTTGCCCGTGTTGCGCGTCTCGGCGACCTTGTTTTCGCAGGGTTGGCTAT 3240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 GGTTGATGGTTGCCCGTGTTGCGCGTCTCGGCGACCTTGTTTTCGCAGGGTTGGCTAT 3240

Qy  3241 GCCGCCAGATGGTGGGCCATTACCGCACCTTTACGCTGCAGTGTCTCGAACGTGG 3300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 GCCGCCAGATGGTGGGCCATTACCGCACCTTTACGCTGCAGTGTCTCGAACGTGG 3300

Qy  3301 CACGCTGTCAGCGATGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC 3360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 CACGCTGTCAGCGATGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC 3360

Qy  3361 TATCTTCAGATTAGGAATCTCTGGCCACTAGCTACATGGGATTTGTTGTGTGACAACGTGTT 3420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 TATCTTCAGATTAGGAATCTCTGGCCACTAGCTACATGGGATTTGTTGTGTGACAACGTGTT 3420

Qy  3421 GTATACTGCTCACCATGGCAGCAAGGGGCGCCGGTTGGCTCATCCCACAGGCTCTATACA 3480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421 GTATACTGCTCACCATGGCAGCAAGGGGCGCCGGTTGGCTCATCCCACAGGCTCTATACA 3480

Qy  3481 CCCAATAACCGTTGACGCGGCTAATGACCAGGACAATCATCAACCACCACCATGTGGAGCTGG 3540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481 CCCAATAACCGTTGACGCGGCTAATGACCAGGACAATCATCAACCACCACCATGTGGAGCTGG 3540

Qy  3541 GTCCCTTACTCGGTGCTCTTGCCGGGAGACCAAGGGTATCTGGTAACGACTGGGGTC 3600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541 GTCCCTTACTCGGTGCTCTTGCCGGGAGACCAAGGGTATCTGGTAACGACTGGGGTC 3600

Qy  3601 ATTGGTGAGGTCAACAAATCCGATGACCCTATTGGTGTGTGTGCTCCTCCCGGGCATGTTATTGG 3660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601 ATTGGTGAGGTCAACAAATCCGATGACCCTATTGGTGTGTGTGCTCCTCCCGGGCATGTTATTGG 3660

Qy  3661 GGCTGTTGCCAAGGGTTCTTCTGGCGGTTCAGTGCAGTCAGATTAGGGTTAGGCCGTT 3720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661 GGCTGTTGCCAAGGGTTCTTCTGGCGGTTCAGTGCAGTCAGATTAGGGTTAGGCCGTT 3720

Qy  3721 GATGTGTGCTGATACAtCCATCCCCCAGTACACAGCACATGCCCACTCTTGATACAAAACCTAC 3780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721 GATGTGTGCTGATACAtCCATCCCCCAGTACACAGCACATGCCCACTCTTGATACAAAACCTAC 3780

Qy  3781 GGTGTGTGCTCGGTGCTCTGTAGATACACATGCCACTCTTGATACAGAAGAACTGCTGACAA 3840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781 GGTGTGTGCTCGGTGCTCTGTAGATACACATGCCACTCTTGATACAGAAGAACTGCTGACAA 3840

Qy  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCCAGCGGCAAGTCAAC 3900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCCAGCGGCAAGTCCAGTGT 3960


Qy  3901 CAAATTACCACCTTCTTACATGCAGAGAAGTATGAGGTCTGGTCCTAAATCCCAGTGTCA 3960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 CAAATTACCACCTTCTTACATGCAGAGAAGTATGAGGTCTGGTCCTAAATCCCAGTGT 3960

Qy  3961 GGCTACAACAGCATCAATGCCAAAGTACATGCGACGCGACGTACGGGCGTACGGGCTGAATCCAAATTG 4020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 GGCTACAACAGCATCAATGCCAAAGTACATGCGACGCGACGTACGGGCGTACGGGCTGAATCCAAATTG 4020

Qy  4021 CTATTTAATGGCAAATGTACCAACACAGGGGCTTCACTACGTAAGACGCACATATGGCAT 4080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 CTATTTAATGGCAAATGTACCAACACAGGGGCTTCACTACGTAAGACGCACATATGGCAT 4080

Qy  4081 GTACCTGACCGGACGATGTTCCCGGAACTATGTAATCATTTGTGACGAATGGCCATGC 4140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 GTACCTGACCGGACGATGTTCCCGGAACTATGTAATCATTTGTGACGAATGGCCATGC 4140

Qy  4141 TACCGATGCAACCACCGTGTTGGGCATTGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA 4200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 TACCGATGCAACCACCGTGTTGGGCATTGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA 4200

Qy  4201 TGTTAGGCTAGTGTTCTTGCCACGGCGTGGCGTGTGGGGACAC 4260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201 TGTTAGGCTAGTGTTCTTGCCACGGCGTGGCGTGTGGGGACAC 4260

Qy  4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAAGAT 4320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAAGAT 4320

Qy  4321 TAAGGAGGAAAAATCTGAAGAAGGAGACACACCTTATCTTTGAGGCTACCACCAAAAAACACTG 4380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321 TAAGGAGGAAAAATCTGAAGAAGGAGACACACCTTATCTTTGAGGCTACCACCAAAAAACACTG 4380

Qy  4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGAATAACAGCTGTCTCTTCTTACTATAGGGG 4440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGAATAACAGCTGTCTCTTCTTACTATAGGGG 4440

Qy  4441 ATGTGACATCTCAAAAATCCCTGAGGGGACTTTGATTGACTTTGTGTAGTGCCACTGATGCCTTGTG 4500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441 ATGTGACATCTCAAAAATCCCTGAGGGGACTTTGATTGACTTTGTGTAGTGCCACTGATGCCTTGTG 4500

Qy  4501 TACAGGGTACACTGGTGACTTGATTCCGTGTATGACTTCATCCCGTAGGAAGGCAC 4560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501 TACAGGGTACACTGGTGACTTGATTCCGTGTATGACTTCATCCCGTAGGAAGGCAC 4560

Qy  4561 ATGCCATGTTGACCTTGACCCTGACCTTCACCATGGGTTCGTTGTGTGGGGTTTCAGC 4620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561 ATGCCATGTTGACCTTGACCCTGACCTTCACCATGGGTTCGTTGTGTGGGGTTTCAGC 4620

Qy  4621 AATAGTTAAAGGCCAGGCTAGGGGCCGTAGGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTTA 4680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621 AATAGTTAAAGGCCAGGCTAGGGGCCGTAGGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTTA 4680

Qy  4681 TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTGTCATCAACAGAGCAAACTATTCTGGAAGCCTT 4740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681 TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTGTCATCAACAGAGCAAACTATTCTGGACAC 4800

Qy  4741 CGACGCGCCACCCAACCTGGGTTACCTGGAGATTAGGGGTCAGTCAGATTAGGGCTGGGCTCTATCGGAGTGGGCTGATCT 4800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741 CGACGCGCCACCCAACCTGGGTTACCTGGAGATTAGGGGTCAGTCAGATTAGGGCTGGGCTCTATCGGAGTGGGCTGATCT 4800

Qy  4801 CTATCGGTCAACCCGACCTTCATTTGTCAACTACGGTCAAACTGTTGTCAAAAGAACTGCTGACAA 4860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801 CTATCGGTCAACCCGACCTTCATTTGTCAACTACGGTCAAACTGTTGTCAAAAGAACTGCTGACAA 4860

Qy  4861 GGTGTGTGCTGATACCATCCCCAGTACCACCTCTTGATACAAACCTACTCATTTGTCAACTACGGTCAATACTGCTGACAA 4920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861 GGTGTGTGCTGATACCATCCCCAGTACCACCTCTTGATACAAACCTACTCATTTGTCAACTACGGTCAATACTGCTGACAA 4920

Qy  4921 TTATGTTTTGTTGACTGCGAGCCCAACTACTGTGTCATCAGTGTGATGTGCTCC 4980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Qy 6061 TTGGCAGTATGTGTGCAATTTCTTTGTGTGATTTGCTTTAATGTCCTTAAAGCTGAGTTCA 6120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6061 TTGGCAGTATGTGTGCAATTTCTTTGTGTGATTTGCTTTAATGTCCTTAAAGCTGAGTTCA 6120

Qy 6121 GAGCATGGTTAACATTCCTGGTGTCCTTTCTACAGCTGCCAGAGGGGTACAAGGGCCC 6180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6121 GAGCATGGTTAACATTCCTGGTGTCCTTTCTACAGCTGCCAGAGGGGTACAAGGGCCC 6180

Qy 6181 CTGGATTGGATCAGGTATGCTCCAAGCACGCTGTCCATGCGGTGCGAACTCATCTTTC 6240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6181 CTGGATTGGATCAGGTATGCTCCAAGCACGCTGTCCATGCGGTGCGAACTCATCTTTC 6240

Qy 6241 TGTTGAGAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG 6300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6241 TGTTGAGAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG 6300

Qy 6301 AGGGGCGTGTTCCAGTCAACGCTAGGCTGTGTGGGCTGGTCGGCTAGACCGAACTGATTG 6360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6301 AGGGGCGTGTTCCAGTCAACGCTAGGCTGTGTGGGCTGGTCGGCTAGACCGAACTGATTG 6360

Qy 6361 GACTAGTCTTGTCGTCAATTATGCGGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA 6420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6361 GACTAGTCTTGTCGTCAATTATGCGGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA 6420

Qy 6421 TCACATTTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCAAC 6480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6421 TCACATTTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCAAC 6480

Qy 6481 CTTGAGAGCTGCAGTGGCCGTGACGGGCTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6481 CTTGAGAGCTGCAGTGGCCGTGACGGGCTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540

Qy 6541 AACTCCTTGGACGACATCCGCTTGCGTGTTCCTGACGGTCCTGCGGTAAGGGTAAAACTGTTAA 6600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6541 AACTCCTTGGACGACATCCGCTTGCGTGTTCCTGACGGTCCTGCGGTAAGGGTAAAACTGTTAA 6600

Qy 6601 GCTTCCCTTCCCGCGTTACGGTCACACACCTGGTGTGCGGCATGCAACTTAATTTGCGTGA 6660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6601 GCTTCCCTTCCCGCGTTACGGTCACACACCTGGTGTGCGGCATGCAACTTAATTTGCGTGA 6660

Qy 6661 TGCACTTGAGACAAATGACTGTAATTCCACAAAACAACACTCCTAGTGATGAAGCCGCAGT 6720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6661 TGCACTTGAGACAAATGACTGTAATTCCACAAAACAACACTCCTAGTGATGAAGCCGCAGT 6720

Qy 6721 GTCCGCTCTTGTTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGCGCAATTTC 6780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6721 GTCCGCTCTTGTTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGCGCAATTTC 6780

Qy 6781 AGCTGGCGGTTGACACCACCAAACTGCCAGCCCCTCACCTGCTCCCCCTCCGAGATCAGCT 6840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6781 AGCTGGCGGTTGACACCACCAAACTGCCAGCCCCTCACCTGCTCCCCCTCCGAGATCAGCT 6840

Qy 6841 GCGCCAGTTCCGGCAAGAACTGGTTCGCTTACCTTGCCTCCGTTCAACGAAGTGACCCGTCCC 6900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6841 GCGCCAGTTCCGGCAAGAACTGGTTCGCTTACCTTGCCTCCGTTCAACGAAGTGACCCGTCCC 6900

Qy 6901 AGGAGTGTCATGTCCTACGCAATTGGATGTGCAATGACGAAGCGGCGAAACAGGCCGGGTGA 6960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6901 AGGAGTGTCATGTCCTACGCAATTGGATGTGCAATGACGAAGCGGCGAAACAGGCCGGGTGA 6960

Qy 6961 CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGCGACAGTCCCTGTTGGGAGCGGGGTGA 7020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6961 CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGCGACAGTCCCTGTTGGGAGCGGGGTGA 7020

Qy 7021 GTGTAACCCTTTCACTGCAATTGGAGTGTCATATCAGCTTGCAT 7080
         ||||||||||||||||||||||||||||||||||||||||||||
Db 7021 GTGTAACCCTTTCACTGCAATTGGAGTGTCATATCAGCTTGCAT 7080

Qy 7081 TTTACCCAGTTACCCTCCCAAAAAGGAGTCTCTGAATGGTCAGAAGTTGGTCGAC 7140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7081 TTTACCCAGTTACCCTCCCAAAAAGGAGTCTCTGAATGGTCAGAAGTTGGTCGAC 7140

Db 4921 TTATGTTTGTTGCTGAGCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTGCTCC 4980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 4921 CAATGACGCACCAGGTGGCAGGGAGCCGGCTTGGGAAAAACCTTGGGGTTCTGTG 5040
Db 4981 CAATGACGCACCAGGTGGCAGGGAGCCGGCTTGGGAAAAACCTTGGGGTTCTGTG 5040

Qy 5041 GCGCTTGGACGGGCGCCTGTCTCCTGGCCCAGAGCGCAGCCCAGCGAGTGACCAGATACCA 5100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5041 GCGCTTGGACGGGCGCCTGTCTCCTGGCCCAGAGCGCAGCCCAGCGAGTGACCAGATACCA 5100

Qy 5101 AATGTGCTTCACTGAAGTCAATACTTCTGGGACAGCAGCCGCACTCGCTGTTGGGGTTGGAGT 5160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5101 AATGTGCTTCACTGAAGTCAATACTTCTGGGACAGCAGCCGCACTCGCTGTTGGGGTTGGAGT 5160

Qy 5161 GGCTATGGCTTATCTAGCCATTGCTGACACTTTGGCGCCGCCACCCTGTGTGCCGGTTGCTGTC 5220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5161 GGCTATGGCTTATCTAGCCATTGCTGACACTTTGGCGCCGCCACCCTGTGTGCCGGTTGCTGTC 5220

Qy 5221 TATTACATCAGTCCCTACACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAGAAATCGT 5280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5221 TATTACATCAGTCCCTACACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAGAAATCGT 5280

Qy 5281 GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGGCCATGTTGCTGCAATTGACAAGCTGAA 5340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5281 GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGGCCATGTTGCTGCAATTGACAAGCTGAA 5340

Qy 5341 GAGTACAATCACCACAACTAGTCCTTTCACAATTGGAAAACCGCACCTTGAAAAACTTAACAC 5400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5341 GAGTACAATCACCACAACTAGTCCTTTCACAATTGGAAAACCGCACCTTGAAAAACTTAACAC 5400

Qy 5401 CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGTGGTTTAGT 5460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5401 CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGTGGTTTAGT 5460

Qy 5461 CACTTTACCTGACAAATCCCTTTGCATCATGCCGTGTTTGCTTTCATTGCGGGTATTACTAC 5520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5461 CACTTTACCTGACAAATCCCTTTGCATCATGCCGTGTTTGCTTTCATTGCGGGTATTACTAC 5520

Qy 5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTTGGAGGCCGCAATTGCGTCCAA 5580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTTGGAGGCCGCAATTGCGTCCAA 5580

Qy 5581 GCTTACAGACGGTAGAGGCGCACTGGGCGTTCATGATGGCCGGGGCTGCGGGAACAGCTCT 5640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5581 GCTTACAGACGGTAGAGGCGCACTGGGCGTTCATGATGGCCGGGGCTGCGGGAACAGCTCT 5640

Qy 5641 TGTGTACATGGACATCGGTGGGTTTGTCTTTGTTGCTTTGCTAGGCGGCTATGCTGCCGCCTC 5700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5641 TGTGTACATGGACATCGGTGGGTTTGTCTTTGTTGCTTTGCTAGGCGGCTATGCTGCCGCCTC 5700

Qy 5701 ATCCACTGCTGTTGACATTTAAAATGCTTGATGGGTGAGTGGGTCAGGTGGCCCACTATGGATCAGCT 5760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5701 ATCCACTGCTGTTGACATTTAAAATGCTTGATGGGTGAGTGGGTCAGGTGGCCCACTATGGATCAGCT 5760

Qy 5761 TGCTGGTTAGTCTACTCCGCGTTCAATCCGGCGTTCATGATGGCCGGGGCTGCGGAGTTGTGTCAGC 5820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5761 TGCTGGTTAGTCTACTCCGCGTTCAATCCGGCGTTCATGATGGCCGGGGCTGCGGAGTTGTGTCAGC 5820

Qy 5821 TTGTGCAATGTTTGCTTTGACAACAGCAGGGCCAGATCACTACTGGCCAACAGACTTCTTAC 5880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5821 TTGTGCAATGTTTGCTTTGACAACAGCAGGGCCAGATCACTACTGGCCAACAGACTTCTTAC 5880

Qy 5881 TATGTTGCTAGGAGCAACACTGGGCATCTATGTAAGTGAATGCCACTACACCGTCGACATCCG 5940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5881 TATGTTGCTAGGAGCAACACTGGGCATCTATGTAAGTGAATGCCACTACACCGTCGACATCCG 5940

Qy 5941 CAGGAAGATACTGGGCATTCTGGAGGCATCACCCCCTGGAGTGTCATATCAGCTTGCAT 6000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5941 CAGGAAGATACTGGGCATTCTGGAGGCATCACCCCCTGGAGTGTCATATCAGCTTGCAT 6000

Qy 6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGGCCCTCATTGCTTGGGGTCTTAGAGAT 6060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGGCCCTCATTGCTTGGGGTCTTAGAGAT 6060

```
Qy  7141 GGCTACAACCGCTTCCAGCTACGTTACTCGGCCCCCGTACCCTAAGATACGGGGAAAGGA 7200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141 GGCTACAACCGCTTCCAGCTACGTTACTCGGCCCCCGTACCCTAAGATACGGGGAAAGGA 7200

Qy  7201 TTCCACTCAGTCAGCCCCGCCAAACGGCCTACACCTGGACGACGTGATTAGCTTCAAAAC 7260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201 TTCCACTCAGTCAGCCCCGCCAAACGGCCTACACCTGGACGACGTGATTAGCTTCAAAAC 7260

Qy  7261 TTCGTGCAGCATGAGCTACACCTGGACCGACCTGGACTCGGACGATTAGCTTCAAACTGC 7320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261 TTCGTGCAGCATGAGCTACACCTGGACCGACCTGGACTCGGACGATTAGCTTCAAACTGC 7320

Qy  7321 TCTGTCTGCAACTCGGGCCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGT 7380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7321 TCTGTCTGCAACTCGGGCCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGT 7380

Qy  7381 GACTGAGCGCGGGGATGCGGAGCTTAGAAAACAAAAGTCACTATTAATAGACAACCTCT 7440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381 GACTGAGCGCGGGGATGCGGAGCTTAGAAAACAAAAGTCACTATTAATAGACAACCTCT 7440

Qy  7441 GTTCCCCCATCATACCACAAGCAAGGATGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441 GTTCCCCCATCATACCACAAGCAAGGATGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500

Qy  7501 CGGGTGTCATGTGGGACTATGATGAACTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTC 7560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7501 CGGGTGTCATGTGGGACTATGATGAACTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTC 7560

Qy  7561 CCACATCACTGGCCTTCGGGGCCACTGAGTGTCGTTCTGGAGCAGCCCCAGGCCCTGTTCT 7620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7561 CCACATCACTGGCCTTCGGGGCCACTGAGTGTCGTTCTGGAGCAGCCCCAGGCCCTGTTCT 7620

Qy  7621 GGACTTGACGAAGTGTGTCGAGGCAGTGAGTCCAGTCCATTATCGGCAAACTGTGAT 7680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7621 GGACTTGACGAAGTGTGTCGAGGCAGTGAGTCCAGTCCATTATCGGCAAACTGTGAT 7680

Qy  7681 AGTTCCAAAGGAGGAGGAGTCTTCGTGAAGACCCCCAGAAACCAACAAAGAAAACCCCCAAG 7740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7681 AGTTCCAAAGGAGGAGGAGTCTTCGTGAAGACCCCCAGAAACCAACAAAGAAAACCCCCAAG 7740

Qy  7741 GCTTATCTCGTACCCCCCACCTTGAAATGAGATGTTGAGAAGATGTACTACGGTCAGGT 7800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7741 GCTTATCTCGTACCCCCCACCTTGAAATGAGATGTTGAGAAGATGTACTACGGTCAGGT 7800

Qy  7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTACGGGTTTGTAGATCCACGTAC 7860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTACGGGTTTGTAGATCCACGTAC 7860

Qy  7861 CCGTGCAAGCGTCGTTGTGTCGATGTGGTCACCCGATGCGAGTCGCAGCCGACATGCGGATAC 7920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7861 CCGTGCAAGCGTCGTTGTGTCGATGTGGTCACCCGATGCGAGTCGCAGCCGACATGCGGATAC 7920

Qy  7921 AGTGTGTTTGACAGTAAAGCTTGATACCATCACACATGGTGGAGACAGACATCTACTC 7980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7921 AGTGTGTTTGACAGTAAAGCTTGATACCATCACACATGGTGGAGACAGACATCTACTC 7980

Qy  7981 AGCAGCTAAACTCAGTGACCAACACCACCGAGTCTGGCATTCACACCATTGCGAGGCAGTATA 8040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7981 AGCAGCTAAACTCAGTGACCAACACCACCGAGTCTGGCATTCACACCATTGCGAGGCAGTATA 8040

Qy  8041 CGCTGGAGGACCGATGATCGCTTATGATGACGGATCCGAGAGATCGGATATCGTAGGTGTAGGTC 8100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8041 CGCTGGAGGACCGATGATCGCTTATGATGACGGATCCGAGAGATCGGATATCGTAGGTGTAGGTC 8100

Qy  8101 TTCCGGCGTCTATACTACCTCAAGTTCCAACAGTTGACCTGGCTGAAGGTAAATGC 8160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8101 TTCCGGCGTCTATACTACCTCAAGTTCCAACAGTTGACCTGGCTGAAGGTAAATGC 8160

Qy  8161 TGCAGCCGAACAGGCTGGCATGGACCCTCGCTTCCTTATTTGCGGCGATGATTGCAC 8220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161 TGCAGCCGAACAGGCTGGCATGGACCCTCGCTTCCTTATTTGCGGCGATGATTGCAC 8220

Qy  8221 CGTAATTTGGAAGAGCGCCGGAGCAGCCGCGGAGCAGATGCAGATGGCAGACAAAACAAGCAATGCCCAAATACAGTTTGCTAG 8280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8221 CGTAATTTGGAAGAGCGCCGGAGCAGCCGCGGAGCAGATGCAGATGGCAGACAAAACAAGCAATGCCCAAATACAGTTTGCTAG 8280

Qy  8281 CTGGATGAAGGTGATGGGTGCGATGGGGTGCACCACCAAGATTGTGTGCCTCAACCCAAATACAGTTTGA 8340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281 CTGGATGAAGGTGATGGGTGCGATGGGGTGCACCACCAAGATTGTGTGCCTCAACCCAAATACAGTTTGA 8340

Qy  8341 AGAATTAACATCATGCTCATCATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8341 AGAATTAACATCATGCTCATCATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400

Qy  8401 CTACTTTCTTACAAGAGATCTCGTATCCCCCTGGGCAGGTGTCTCGCGAGGGTCTGGG 8460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8401 CTACTTTCTTACAAGAGATCTCGTATCCCCCTGGGCAGGTGTCTCGCGAGGGTCTGGG 8460

Qy  8461 ATACAACCCAGTGCTGCGTGGATTGGGTATCTAATACATCACTCAATACATCACTCACCATGTTGTGGGT 8520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8461 ATACAACCCAGTGCTGCGTGGATTGGGTATCTAATACATCACTCAATACATCACTCACCATGTTGTGGGT 8520

Qy  8521 TAGCCGTGTGTTGGCTGTCGTCCATTCATGCAGAGATGCTCTTTGAGACAACAAACTTCCCGA 8580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8521 TAGCCGTGTGTTGGCTGTCGTCCATTCATGCAGAGATGCTCTTTGAGACAACAAACTTCCCGA 8580

Qy  8581 GACTGTGACCTTTGACTGTGATATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8581 GACTGTGACCTTTGACTGTGATATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640

Qy  8641 CATCATTGCTGCGTGTGCACGGTATTGAGGGCTTTCTCCGGTGTGTGCGCTACACCAACGCTGA 8700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8641 CATCATTGCTGCGTGTGCACGGTATTGAGGGCTTTCTCCGGTGTGTGCGCTACACCAACGCTGA 8700

Qy  8701 GATCCTCAGAGTTCCCAATCACTAACAGACATGACCATGCCCCCCTGCGAGCCTGGCG 8760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8701 GATCCTCAGAGTTCCCAATCACTAACAGACATGACCATGCCCCCCTGCGAGCCTGGCG 8760

Qy  8761 AAAGAAAGCCAGGGCGGTCCTGCCAAGAGGCGTGGCGGAGCACACGCCAAAATT 8820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8761 AAAGAAAGCCAGGGCGGTCCTGCCAAGAGGCGTGGCGGAGCACACGCCAAAATT 8820

Qy  8821 GGCTCGCTTCCTTCTCTGCGATGCCATCTCTAGACCTTCTACCAGATTTGGATAAGAGACGAG 8880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8821 GGCTCGCTTCCTTCTCTGCGATGCCATCTCTAGACCTTCTACCAGATTTGGATAAGAGACGAG 8880

Qy  8881 CGTGGCTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGGGGATGTGTT 8940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8881 CGTGGCTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGGGGATGTGTT 8940

Qy  8941 TATTACACCACAGAGAAGATTGCGAGAGTTCCTGTGAAGTATTTGGCTGTCATTGTTTT 9000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8941 TATTACACCACAGAGAAGATTGCGAGAGTTCCTGTGAAGTATTTGGCTGTCATTGTTTT 9000

Qy  9001 TGCCCTAGGGCTCATTGCTGTGTTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9001 TGCCCTAGGGCTCATTGCTGTGTTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060

Qy  9061 CTAACAGTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGGCAACAGCCCCGGGC 9120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9061 CTAACAGTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGGCAACAGCCCCGGGC 9120

Qy  9121 TTAACGACCCCGC 9133
         |||||||||||||
Db  9121 TTAACGACCCCGC 9133
```

REFERENCE   Unclassified.
             1 (bases 1 to 9143)
AUTHORS   Simons,J.N., Desai,S.M. and Mushahwar,I.K.
TITLE   Reagents and methods useful for controlling the translation of
             hepatitis GBV proteins
JOURNAL   Patent: US 5955318-A 32 21-SEP-1999;
FEATURES   Location/Qualifiers
source   1..9143
             /organism="unknown"
             /mol_type="unassigned DNA"
ORIGIN

Query Match   96.4%; Score 9059.8; DB 6; Length 9143;
Best Local Similarity   99.6%; Pred. No. 0;
Matches 9096; Conservative   0; Mismatches   37; Indels   4; Gaps   1;

```
Qy       1 ACCACAAACACTCCAGTTTGTTACACTTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ACCACAAACACTCCAGTTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCCTAG   60

Qy      61 CAGGGGCTGGGGGATTTCCCCTGCCCGTCTGCAGAAGGGTGAGCCAACCACCTTAGTAT  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CAGGGGCGTGGGGGATTTCCCCTGCCCGTCTGCAGAAGGGTGAGCCAACCACCTTAGTAT  120

Qy     121 GTAGGGCGGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GTAGGGCGGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180

Qy     181 CCTGATGGGCGTTCATGGTTCGGTGGTGTGGCGCGCTTTAGGCAGCCTCCAGCCCACCA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CCTGATGGGCGTTCATGGTTCGGTGGTGTGGCGCGCTTTAGGCAGCCTCCACGCCCCACCA  240

Qy     241 CCTCCCAGATAGAGCGGCGGGACGCTGTAGGGAAGACCGGGACCGGGACCCACTACCAAGGACG  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CCTCCCAGATAGAGCGGCGGGACGCTGTAGGGAAGAGCGGGACCGGGACCCACTACCAAGGACG  300

Qy     301 CAGACCTCTTTTGAGTATCACGGCCTCCGAAGTAGTTGGGCAAGCCACCTATATGTGT  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CAGACCTCTTTTGAGTATCACGGCCTCCGAAGTAGTTGGGCAAGCCACCTATATGTGT  360

Qy     361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT  420

Qy     421 CTGGGAGTCTCGTAGACCGTAGACACATGCCGTGTTATTTCTACTCACAAGTCCTGTACC  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CTGGGAGTCTCGTAGACCGTAGACACATGCCGTGTTATTTCTACTCACAAACAAGTCCTGTACC  480

Qy     481 TGCGCCCAGAACGCGCCAAGACGCCAAGAGCAGACGCAGGCTTCATATCCTGTCCATTAAAAC  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGCGCCCAGAACGCGCCAAGACGCCAAGAGCAGACGCAGGCTTCATATCCTGTCCATTAAAAC  540

Qy     541 ATCTGTTGAAAGGGGACAACGACGAGCAAAGCCAAGTTCAGCGCGATGCTCGGCCTCGTAA  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 ATCTGTTGAAAGGGGACAACGACGAGCAAAGCCAAGTTCAGCGCGATGCTCGGCCTCGTAA  600

Qy     601 TTTACAAAATTGCTGTATCCATGATGGCTTGCAGACATTGGCTCGAGGCTGCTTGCCAGC  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TTTACAAAATTGCTGTATCCATGATGGCTTGCAGACATTGGCTCGAGGCTGCTTTGCCAGC  660

Qy     661 TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTCCGCAATCTTGGAATCCTTCTGA  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 TGCGCCATCGGTGGGGACGCCAAGACCCTCGCCATAAGTCTCGCAATCTTGGAATCCTTCTGA  720

Qy     721 TTTACCCTTTGGGGTGGATTGGATTGGTGATGTTACAACTCACAACACCTCTAGTAGGCCCGGCTGCT  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TTTACCCTTTGGGGTGGATTGGATTGGTGATGTTACAACTCACAACACCTCTAGTAGGCCCGGCTGCT  780

Qy     781 GGCAGGAGCGGCGGTCGTTCGACCAGTCTGCCGAGATAGTACGCTTCGCTGGGATGGAGTGGAGTGCAA  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GGCAGGAGCGGCGGTCGTTCGACCAGTCTGCCGAGATAGTACGCTTCGCTGGGATGGAGTGGAGTGCAA  840

Qy     841 CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTTGTGGTATGTCTGCTATCTTTGGGCCTG  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTTGTGGTATGTCTGCTATCTTTGGGCCTG  900

Qy     901 TCCCTGTAGTGGGGCGGCCGGGGTCACTGACCACCAGACACAAATACCACAATCCTGACCAATTG  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TCCCTGTAGTGGGGCGGCCGGGGTCACTGACCACCAGACACAAATACCACAATCCTGACCAATTG  960

Qy     961 CTGCCAGGTAATCAGTTATCTATTGTCTCCTTCCACTTGCCTACACGAGCCTGGTTG  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 CTGCCAGGTAATCAGTTATCTATTGTCTCCTTCCACTTGCCTACACGAGCCTGGTTG  1020

Qy    1021 TGTGATCTGTGCGACGAGTGCTGGGTTCCCGCCAATCCTCACACCCTTCCAA  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 TGTGATCTGCGCGACGAGTGCTGGGTTCCCGCCAATCCGTACAATCTCACACCCTTCCAA  1080

Qy    1081 TTGGACTGGCACGGACTCCTTCTTGGCTGACCACCATTGATTTTGTTATGGGGCTCTTGT  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 TTGGACTGGCACGGACTCCTTCTTGGCTGACCACCATTGATTTTGTTATGGGCCGCTTCTTGT  1140

Qy    1141 GACCTGTGACGCCCTGACATTGGTGAGTTGGGTGTGTGGCGTGTATTAGTCGGTGACTG  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 GACCTGTGACGCCCTGACATTGGTGAGTTGGGTGTGTGGCGTGTATTAGTCGGTGACTG  1200

Qy    1201 GCTTGTCAGGCACTGGCTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGA  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 GCTTGTCAGGCACTGGCTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGA  1260

Qy    1261 AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT  1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT  1320

Qy    1321 CGAGGCTGTCATCTTCTTGACCAAACTGCTTCACAAGTACCATACGTATTGCGACTAT  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 CGAGGCTGTCATCTTCTTGACCAAACTGCTTCACAAGTACCATACGTATTGCGACTAT  1380

Qy    1381 GTTTAGCAGTGTACACTACCTGGCGGTTGGCGCTCTGATCTACTACATGCCTCGGGGCAA  1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 GTTTAGCAGTGTACACTACCTGGCGGTTGGCGCTCTGATCTACTACATGCCTCGGGGCAA  1440

Qy    1441 GTGGTATCAGTGCTTATGCTTTACATAGAAGCGACCTCTGGAAACCCCAT  1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 GTGGTATCAGTGCTTATGCTTTACATAGAAGCGACCTCTGGAAAACCCCTAT  1500

Qy    1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTTGCTCGCCTTGATGATACCATGTCC  1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTTGCTCGCCTTTGATGATACCATGTCC  1560

Qy    1561 TTGCCACTCTTATTTGAGTGAGAATGTCAGAGTCATTGTTACAGTCCAAAGTGGAC  1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAGTCATTGTTACAGTCCAAAGTGGAC  1620

Qy    1621 CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGTGTACCCCTATACAATCCCTG  1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 CAGGCCTGTCACTCTAGAGTATAACAACTCCATATCTTGTGTACCCCTATACAATCCCTG  1680

Qy    1681 TGCGAGGGGATGTATGGTAAATTCAAAAATACACATGGGTTGCTCGCCTATTCGCAA  1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 TGCGAGGGGATGTATGGTAAATTCAAAAATAACACATGGGTTGCTCGCCTATTCGCAA  1740

Qy    1741 TGTGCCATCGTACTGCACTATGGCACGTGAACGACACTCGCAACACTTA  1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 TGTGCCATCGTACTGCACTATGGCACTGTGAACGACACTCGCAACACTTA  1800

Qy    1801 CGAAGCATGCGGTGTAACACCATGGCTACACCGGCCACAACCGGCTCAGCCCTGAA  1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 CGAAGCATGCGGTGTAACACCATGGCTACACCGGCCACAACCGGCTCAGCCCTGA  1860

Qy    1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC  1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC  1920

Qy    1921 AGGCCATTTGTATTTTGAGGGATCAGATACCCTATAGTTTACTTTTATGTGACCCTGTGAA  1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1921  AGGCCATTTGTATTTGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAA  1980

Qy   1981  TTCCACTCTCCTACCCGGAGAGGTTGGGCTAGGTTGCCCGTACCCCACCTGTGGTACG  2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981  TTCCACTCTCCTACCCGGAGAGGTGGGCTAGGTTGCCCGTACCCCACCTGTGGTACG  2040

Qy   2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAGG  2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAGG  2100

Qy   2101  ATTGATCACCAAAGACAAAGCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC  2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101  ATTGATCACCAAAGACAAAGCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC  2160

Qy   2161  TTTGTCTTACGGGAGTTACCACCAAGGCCGTGGTCTAATTCTGTTGGTGGGTGTGTGTGG  2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161  TTTGTCTTACGGGAGTTACCACCAAGGCCGTGGTCTAATTCTGTTGGTGGGTTGTGTGG  2220

Qy   2221  CAGCAAGTATCTTATTTAGCCTACCTCTGTACTTGTCCCTTTGTTTGGGCGGCGCTTC  2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221  CAGCAAGTATCTTATTTAGCCTACCTCTGTACTTGTCCCTTTGTTTGGGCGGCGCTTC  2280

Qy   2281  TGGTTACCCTTTGCGTCTGTCTCCCCAGTCGTATCTCCAGTCGTATCTCCAGCTGGTGGGATGT  2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281  TGGTTACCCTTTGCGTCTGTCTCCCCAGTCGTATCTCCAGTCGTATCTCCAGCTGGTGGGATGT  2340

Qy   2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTTGCTTTGATTTCTTCATCTGTTGCTATCTCCG  2400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTTGCTTTGATTTCTTCATCTGTTGCTATCTCCG  2400

Qy   2401  CTGCAGGCTACGTTATGCGTCGCCCTTTTAGGGTTGTGCCACCAGTTATGACTGTGGGCGACTGCT  2460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401  CTGCAGGCTACGTTATGCGTCGCCCTTTTAGGGTTGTGCCACCAGTTATGACTGTGGGCGACTGCT  2460

Qy   2461  AACTTTCTTTGTTGTGTGCAGCAGCTGCTGCCCACCGGCTACCGGTAACCGGTGGTTCACGT  2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461  AACTTTCTTTGTTGTGTGCAGCAGCTGCTGCCCACCGGCTACCGGTAACCGGTGGTTCACGT  2520

Qy   2521  AGTGGCAGGGTTAGTTTGTTGTGGGCCGGCCGTAACCGGTGGTGACCCGGTATAGTCTCTGTGT  2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521  AGTGGCAGGGTTAGTTTGTTGTGGGCCGGCCGTAACCGGTGGTGACCCGGTATAGTCTCTGTGT  2580

Qy   2581  AGGTTCCTTGGCCTCTGGTAGGGCTTTAACCCTTCTTGCATTTGGTTACGCGCTGCTTCAGC  2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581  AGGTTCCTTGGCCTCTGGTAGGGCTTTAACCCTTCTTGCATTTGGTTACGCGCTGCTTCAGC  2640

Qy   2641  TTTTGATACCGGAGATAATTGAGAGGGCTGACATACCACCTGTAGTAGCATTAGTTGTCAT  2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641  TTTTGACACCGGAGATAATTGAGAGGGCTGACATACCACCTGTGCTGGTGCTTAGTTGTCAT  2700

Qy   2701  GTCTCGTTTGCTTTGGCTTCTCCTATGTTTAACATCCAGTCGTTCTTGGGACTGACTCTAGGGT  2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701  GTCTCGTTTGCTTTGGCTTCTCCTATGTTTAACATCCAGTCGTTCTTGGGACTGACTCTAGGGT  2760

Qy   2761  TTGGCAACGTTGGGGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTTCCT  2820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761  TTGGCAACGTTGGGGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTTCCT  2820

Qy   2821  TGTGTGGTTGTTTGTTTCCCCGGTGCCAGCATATGACGCGCGTGGTGACTTCTGTGTGTCA  2880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821  TGTGTGGTTGTTTGTTTCCCCGGTGCCAGCATATGACGCGCGTGGTGACTTCTGTGTGTCA  2880

Qy   2881  CGTAGCTCTTCTATGTTTAACATCCAGTCGTGCAGCATCGTTCTTTGGGACTCTAGGGT  2940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881  CGTAGCTCTTCTATGTTTAACATCCAGTGCAGCATCGTTCTTTGGGACTCTAGGGT  2940

Qy   2941  TAGGGCCCATAGAATGTTGGTGGTCTTGGTTTTGGTTATTCTCATTATGT  3000
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941  TAGGGCCCATAGAATGTTTAATGGCAAATGTACCAACACAGGGGTATTCTCATTATGT  3000

Qy   3001  TCTTAAGTTTTTCCCTCTTAGTGTTTTTGGTTGGTGAAATGGTGTGTTTTTTCTATAAGCACTTGCA  3060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001  TCTTAAGTTTTTCCCTCTTAGTGTTTTCCCGTGAGAATGGTGTGTTTTTTCTATAAGCACTTGCA  3060
```

```
Qy   3061  TGGTGATGTCTTGCCTAATGATTTTGCTCGAACTACCACCATTGCAAGAGCCATTTTTCCC  3120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061  TGGTGATGTCTTGCCTAATGATTTTGCTCGAACTACCACCATTGCAAGAGCCATTTTTCCC  3120

Qy   3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAGAACACGTTGGCGTGTGGGGACAC  3180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAGAACACGTTGGCGTGTGGGGACAC  3180

Qy   3181  GGTTGATGGTTGCCCGGTTGTCCGGCGACCTTGTTTTCGCAGGGTTGGCTAT  3240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181  GGTTGATGGTTGCCCGGTTGTCCGGCGACCTTGTTTTCGGAGGGTTAGCTAT  3240

Qy   3241  GCCGCCAGATGGGTGGGCCATTACCGGCACCTTTACGCTGCAGTGTCTCTGAACGTGGAA  3300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241  GCCGCCAGATGGGTGGGCCATTACCGGCACCTTTACGCTGCAGTGTCTCTGAACGTGGG  3300

Qy   3301  CACGCTGTCAGCGATGGCAGTGCTCTGGCCACTAGCTACATGGCTACATGGGAACTGGAAC  3360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301  CACGCTGTCAGCGATGGCAGTGCTCTGGCCACTAGCTACATGGCTACATGGGACTGGAAC  3360

Qy   3361  TATCTTCAGATTAGGATCTCTGCCACTAGCTACATGGGATTTGTTTGTGACAACGTGTT  3420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361  TATCTTCAGATTAGGATCTCTGCCACTAGCTACATGGGATTTGTTTGTGACAACGTGTT  3420

Qy   3421  GTATACTGCTCACCATGGCAGCAAGGGCGCGCCGGTTGGCTAATGACCGGCTAATTATCCCACCAGGCTCTATACA  3480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421  GTATACTGCTCACCATGGCAGCAAGGGCGCGCCGGTTGGCTAATGACCGGCTAATTATCCCACCAGGCTCCATACA  3480

Qy   3481  CCCAATAACCGTTGACGGGCTAATGACCGGAGACATCTATCAACCACCATGTGGAGCTGG  3540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481  CCCAATAACCGTTGACGGGCTAATGACCGGAGACATCTATCAACCACCATGTGGAGCTGG  3540

Qy   3541  GTCCCTTACTCGGTCGTCTCTTGCGGGGAGACCAAGGGTATCGTGTGGGGGTC  3600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541  GTCCCTTACTCGGTCGTCTCTTGCGGGGAGACCAAGGGTATCGTGTGGGGGGTC  3600

Qy   3601  ATTGGTTGAGGTCAACAAATCCGATGACCCTATTGGTGTGTGCCCTTCCCAT  3660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601  ATTGGTTGAGGTCAACAAATCCGATGACCCTATTGGTGTGTGCCCTTCCCAT  3660

Qy   3661  GGCTGTTGCCAAGGGTTCTTCAGGGTGTACGCGCGATTCGTGCTCTCCTCGGGCATGTTATTGG  3720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661  GGCTGTTGCCAAGGGTTCTTCAGGGTGTACGCGCGATTCGTGCTCTCCTCGGGCATGTTATTGG  3720

Qy   3721  GATGTTCACCGCGCTGCTAGAAATTCTGGCGGTTCAGTCAGCACAGCACATGCCACTCTTGATACAAAACCTAC  3780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721  GATGTTCACCGCGCTGCTAGAAATTCTGGCGGTTCAGTCAGCACAGCACATGCCACTCTTGATACAAAACCTAC  3780

Qy   3781  GGTGTGTGCTGAATACCATCCCAGTACACAGCACACATGGCAGCAAGTCAAC  3840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781  GGTGTGTGCTGAATACCATCCCAGTACACAGCACACATGGCAGCAAGTCAAC  3840

Qy   3841  TGTGCCTAACGAGTATTCAGTGCAAATTTTAATTGCCCCACTGGCAGCGGCAAGTCCCAGTGT  3900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841  TGTGCCTAACGAGTATTCAGTGCAAATTTTAATTGCCCCACTGGCAGCGGCAAGTCCCAGTGT  3900

Qy   3901  CAAATTACCACCTTCTTACATGCAGGAGAAGGATGAGGTCTTGGTCTTCCTAAATCCCAGTGT  3960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901  CAAATTACCACCTTCTTACATGCAGGAGAAGGATGAGGTCTTGGTCTTCCTAAATCCCAGTGT  3960

Qy   3961  GGCTACAACAGCATCAATGCCAAAGTACATGCACGCGGCTACGGCCGTGAATCAAATTG  4020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961  GGCTACAACAGCATCAATGCCAAAGTACATGCACGCGGCTACGGCCGTGAATCAAATTG  4020

Qy   4021  CTATTTTAATGGCAAATGTACCAACACATGGCACTACGCACACATATGGCAT  4080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021  CTATTTTAATGGCAAATGTACCAACACACAGGCACTTACTACGCACACATATGGCAT  4080

Qy   4081  GTACCTGACCGGACGGCATGTTCCCGGAACTATGATGTAATCATTTGTGACGCAATGCCATGC  4140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081  GTACCTGACCGGACGGCATGTTCCCGGAACTATGATGTCATCATTTGTGACGAATGCCATGC  4140
```

Qy   4141 TACCGATGCAACCACCGTGTTGGGCATTGGAAAGGTCCTAACCGAAGCTCATCCAAAAA 4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 TACCGATGCAACCACCGTGTTGGGCATTGGAAAGGTTCTAACCGAAGCTCCATCCAAAAA 4200

Qy   4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCCTGGAGTAATCCCTGGAGTAATCCC 4260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCCTGGAGTAATCCCTGGAGTAATCATGC 4260

Qy   4261 CAACATAACTGAGATTCAATTAACCACACTAGTCTTCACATTGGAAAAAAGAT 4320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 CAACATAACTGAGATTCAATTAACCACACTAGTCTTCACATTGGAAAAAACTGAAAAAGAT 4320

Qy   4321 TAAGGAGGAAAATCTGAAGAAAGGGAGACACCTTATCTTGAGGCTACCAAAAAACTG 4380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321 TAAGGAGGAAAATCTGAAGAAAGGGAGACACCTTATCTTGAGGCTACCAAAAAACTG 4380

Qy   4381 TGATGAGCTTGCTAACCGAGTTAGCTGAAATAACAGCTGTCTCTTACTATAGGGG 4440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 TGATGAGCTTGCTAACCGAGTTAGCTGAAATAACAGCTGTCTCTTACTATAGGGG 4440

Qy   4441 ATGTGACATCTCAAAAATCCCTGAGGGGCGACTGTGTAGTAGTTGCCACTGTGTG 4500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 ATGTGACATCTCAAAAATCCCTGAGGGGCGACTGTGTAGTAGTTGCCACTGTGCCTTGTG 4500

Qy   4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTGCAGCCTCATGGTAGAGGCAC 4560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTGCAGCCTCATGGTAGAAGGCAC 4560

Qy   4561 ATGCCATGTTGACCCTACTTTCACCATGGGTTCGTGTGTGCGGGTTTCAGC 4620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 ATGCCATGTTGACCCTACTTTCACCATGGGTTCGTGTGTGCGGGTTCTCAGC 4620

Qy   4621 AATAGTTAAAGGCCAGCGTAGGGGCCGTGGGAGAGCCGTGGAGAGCCGTGCATATACTACTA 4680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4621 AATAGTTAAAGGCCAGCGTAGGGGCCGTGGGAGAGCCGTGGAGAGCCGTGCATATACTACTA 4680

Qy   4681 TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4681 TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740

Qy   4741 CGACGCAGCCAAGGCCATGGTATGGTTGTCATCAACAGAAGCTCAAACTATTCTGGACAC 4800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 CGACGCAGCCAAGGCCATGGTATGGTTGTCATCAACAGAAGCTCAAACTATTCTGGACAC 4800

Qy   4801 CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT 4860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT 4860

Qy   4861 CTTTTCTATGGTCAACCCCGAACCTTCATTTGTCAATACTACAACTGTGCATCAGTGTGTGACAA 4920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861 CTTTTCTATGGTCAACCCCGAACCTTCATTTGTCAATACTACAACTGTGCATCAGTGTGTGACAA 4920

Qy   4921 TTATGTTTGTGACTGACGGTGACTGGGAGCCCGGCTTGGGCCTTGGGGTTCTGTG 4980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 TTATGTTTGTGACTGACGGTGACTGGGAGCCCGGCTTGGGCCTTGGGGTTCTGTG 4980

Qy   4981 CAATGACGCACCACGGTGGAGGAGCCCGGCTTGCTCCTGGCCAGCGCCAGAGCGTGACCAGATACCA 5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 CAATGACGCACCACGGTGGAGGAGCCCGGCTTGCTCCTGGCCAGCGCCAGAGCGTGACCAGATACCA 5040

Qy   5041 GCGCCTTGACGCGCCGCTGACGCGCCCGGCTTGCTCCTGGCCCAGAGCGTGACTCGCTCGGGTTGGGAGT 5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041 GCGCCTTGACGCGCCGCTGACGCGCCCGGCTTGCTCCTGGCCCAGAGCGTGACTCGCTCGGGTTGGGAGT 5100

Qy   5101 AATGTGCTTCACTGAAGTCAATATCTTCTGGGACAGCACTCGCCACTCGCGTTTGGGCGTTGCTCGTC 5160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101 AATGTGCTTCACTGAAGTCAATATCTTCTGGGACAGCACTCCAAGCACTCGCGTTGGGCGTTGCTCGTC 5160

Qy   5161 GGCTATGCTTATCTAGCCATTGACACTTTTGGCCGCCATTGACACTTTTGGCCGCCGTTGCTCGTC 5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161 GGCTATGCTTATCTAGCCATTGACACTTTTGGCCGCCATTGACACTTTTGGCCGCCGTTGCTCGTC 5220

Qy   5221 TATTACATCAGTCCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAAGAAATCGT 5280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221 TATTACATCAGTCCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAAGAAATCGT 5280

Qy   5281 GGAGGAGTGTGCATCATTCATTCCCTGGAGGCCATGTGTTGCTGCAATGACAAGCTGAA 5340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5281 GGAGGAGTGTGCATCATTCATTCCCTGGAGGCCATGTGTTGCTGCAATCGATAAGCTGAA 5340

Qy   5341 GAGTACAATCACCACAACTAGTCCTTCACATTGGAAACCGCCCTGAAAAAACTTAACAC 5400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5341 GAGTACAATAACCACAACTAGTCCTTCACATTGGAAACCGCCCTGAAAAAACTTAACAC 5400

Qy   5401 CTTTCTTGGCCTCATGCAGCTACAATCCTTGCTATCATAGATTGTGTGGTTTAGT 5460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5401 CTTTCTTGGCCTCATGCAGCTACAATCCTTGCTATCATAGATTGTGTGGTTTAGT 5460

Qy   5461 CACTTTACCTGACAATCCCTTGCATCATGCGTGTTTGCTTTCATTGCGGGTATTACTAC 5520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5461 CACTTTACCTGACAATCCCTTGCATCATGCGTGTTTGCTTTCATTGCGGGTATTACTAC 5520

Qy   5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTGGAGGCGCAATTGCGTCCAA 5580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTGGAGGCGCAATTGCGTCCAA 5580

Qy   5581 GCTTACAGACGCTAGAGGCCACTGGCGTCATGATGGCGGGCGTGCGCGGAACAGCTCT 5640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5581 GCTTACAGACGCTAGAGGCCACTGGCGTCATGATGGCGGGCGTGCGCGGAACAGCTCT 5640

Qy   5641 TGGTACATGGACATCGGTGGGTTTGTCTTTGACAATGCTCGCCGCCTC 5700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5641 TGGTGCAATGTTTGCTTGCAATGTTTGTCTTTGACAACAGCAGGGCCCAACAGAACTTCTTAC 5700

Qy   5701 ATCCACTGCTTGCTTGACATTAAATGCTTGATGGGTGAGTGCCACTGTGACATCCG 5760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5701 ATCCACTGCTTGCTTGACATTAAATGCTTGATGGGTGAGTGCCACTGTGACATCCG 5760

Qy   5761 TGCTGGTTTTAGTCTACTCCGGTCAATCCGGCCGCGCCAGAGTTGTGGGCGTCTTGTGTCAGC 5820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5761 TGCTGGTTTTAGTCTACTCCGGTCAATCCGGCCGCGCCAGAGTTGTGGGCGTCTTGTGTCAGC 5820

Qy   5821 TTGTGCAATGTTTGCTTGACAACAGCAGGGCCCAACAGAACTCACTGGCCACTCGTGACATCCG 5880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5821 TTGTGCAATGTTTGCTTGACAACAGCAGGGCCCAACAGAACTCACTGGCCACTCGTGACATCCG 5880

Qy   5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940

Qy   5941 CAGGAAGATACTGGGCATTCGGAGGCATCTACCCCCTGAGTGTCATATCAGCTGTGCAT 6000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5941 CAGGAAGATACTGGGCATTCGGAGGCATCTACCCCCTGAGTGTCATATCAGCTGTGCAT 6000

Qy   6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCCGGCCTCATTGCTTGCTTAGAGAT 6060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCCGGCCTCATTGCTTGCTTAGAGAT 6060

Qy   6061 TTGGCAGTATGTGTGCAATTTCTTGTGGTTGTCCTTTCTGTGATTTGTGTAAGTCCTTAAAGTCCTTCA 6120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6061 TTGGCAGTATGTGTGCAATTTCTTGTGGTTGTCCTTTCTACAGCTGCAATTGTCCTTTAAAGTCCTGAGTTCA 6120

Qy   6121 GAGCATGGTTAACATTCCTGGTTGTCCTTCTACAGCGCTGCCATGCGGTTGCACAAGGGCCC 6180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6121 GAGCATGGTTAACATTCCTGGTTGTCCTTCTACAGCGCTGCCATGCGGTTGCACAAGGGCCC 6180

Qy   6181 CTGGATTGGATCAGTCAATGCTCCAAGCACGCGTGTCCATGCGGTTGCTGAACTCATCTTTTC 6240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6181 CTGGATTGGATCAGTCAATGCTCCAAGCACGCGTGTCCATGCGGTTGCTGAACTCATCATCTTTTC 6240

Qy   6241 TGTTGAGAATGGTTTGCAAAACTTTACAAGGACCCAGAACTTGTTCAAATTACTGGAG 6300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6241 TGTTGAGAATGGTTTTGCAAAACTTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG 6300

Qy   6301 AGGGGCTGTTCCAGTCAACGCGTAGCCTGTGTGGGCGGCGGTAGACCCAACTGATTG 6360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   6301 AGGGGCTGTTCCAGTCAACGCTAGGCTGTGTGGTCGGCTAGACCGACCCAACTGATTG 6360

Qy   6361 GACTAGTCTTGTCGTCAATTATGGGCGTTAGGGACTACTGTAAATATGAGAAATGGGAGA 6420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6361 GACTAGTCTTGTCGTCAATTATGGGCGTTAGGGACTACTGTAAATATGAGAAATGGGAGA 6420

Qy   6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCCAAC 6480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCCAAC 6480

Qy   6481 CTTGAGAGCTTGCCAGTGCCGTTGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6481 CTTGAGAGCTTGCCAGTGCCGTTGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540

Qy   6541 AACTCCCTGGACACATCTGCTTGCTGTTACGGTCCGGACGGCGTAAGGGTAAAACTGTTAA 6600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6541 AACTCCCTGGACACATCTGCTTGCTGTTACGGTCCGGACGGCGTAAGGGTAAAACTGTTAA 6600

Qy   6601 GCTTCCCTTCCGCGGTTGACGGTCACACACCTGGTGTGCGCATGCACTTAATTTGCGTGA 6660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6601 GCTTCCCTTCCGCGGTTGACGGTCACACACCTGGTGTGCGCATGCACTTAATTTGCGTGA 6660

Qy   6661 TGCACTTGAGACACAAATGACTGTAATTCCACAAACAACACTCCAGTAGGTAAGGAAGCCGCAGT 6720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6661 TGCACTTGAGACACAAATGACTGTAATTCCATAAAACAACACTCCAGTGTGATGATGAAGCCGCAGT 6720

Qy   6721 GTCCGGCTCTGTTTCAAACAGCAGTGCGGGCTACAACCAATTGCTTGAGGCAATTTC 6780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6721 GTCCGGCTCTTGTTTGTTTCAAACAGGCTCTTGTGCGGCGCTACAACCAATTGCTTGAGGCAATTTC 6780

Qy   6781 AGCTGGCGGTTGACACACCACCAAACTGCCAGCCCCCTCCATCGAAGAGGTAGTGGTAAGAAA 6840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6781 AGCTGGGGGTTGACACACCACCAAACTGCCAGCCCCCTCCATCGAAGAGGTAGTGGTAAGAAA 6840

Qy   6841 GCGCCCAGTTCCGGGCAAGAACTGGTTGTTCGCTTACCTTGCCCTGCCCTCCTCCCGAGATCCGTCCC 6900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6841 GCGCCCAGTTCCGGGCAAGAACTGGTTGTTCGCTTACCTTGCCCTGCCCTCCTCCCCGAGATCCGTCCC 6900

Qy   6901 AGGAGTGTCATGTCCTGAAAGCTGCCGACGAAGTGAGTCGCCGGTTAGAGGTTCCTCAAACCT 6960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6901 AGGAGTGTCATGTCCTGAAAGCTGCCGACGAAGTGAGTCGCCGGTTAGAGGTTCCTCAAACCT 6960

Qy   6961 CCCTTCCTTCACCACTGTTCTACAGTTGGCCATGCGCCCTGTTGGGAGCGGGGTGA 7020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6961 CCCTTCCTTCACCACTGTTCTACAGTTGGCCATGCGCCCTGTTGGGAGCGGGGTGA 7020

Qy   7021 GTGTAACCCTTCACTGCTCAATTGGATGTGCAATGACCGAAACAGGCGGCGCCCTGATGA 7080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7021 GTGTAACCCTTTCACTGCTCAATTGGATGTGCAATGACCGAAACAGGCGGCGCCCTGATGA 7080

Qy   7081 TTTACCCAGTTACCCTCCCAAAAAAGAGGAGGTCTTCTGAATGGTTCCTCAAACTGCTTCAAAGT 7140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7081 TTTACCCAGTTACCCTCCCAAAAAAGAGGAGGTCTTCTGAATGGTTCCTCAAACTGCTTCAAAGT 7140

Qy   7141 GGCTACAACCGCTTCAGCTACGTTACTGGCCCCCCGTACCCTAAGGATATACGGGGAAAGGA 7200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7141 GACTACAACCGCTTCAGCTACGTTACTGGCCCCCCGTACCCTAAGGATATACGGGGAAAGGA 7200

Qy   7201 TTCCACTCAGTCAGCCCCCGCCAAACGGCCTACAAAAAGAAGAGTTGGGAAAGAGTGAGTT 7260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7201 TTCCACTCAGTCAGCCCCCGCCAAACGGCCTACAAAAAGAAGAGTTGGGAAAGAGTGAGTT 7260

Qy   7261 TTCGTGCAGCATGAGCTGCAACTCGGACCACGCTGATTAGCTTCAAAACTGCTTCTAAAGT 7320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7261 TTCGTGCAGCATGAGCTGCAACTCGGACCACGCTGATTAGCTTCAAAACTGCTTCTAAAGT 7320

Qy   7321 TCTGTCTGCAACTCGGGCCATCACTTGGCCATCACTTGGTTCCTCAAACAAAGATCATTGGTGTATGT 7380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7321 TCTGTCTGCAACTCGGGCCATCACTTGGCCATCACTTGGTTCCTCAAACAAAGATCATTGGTGTATGT 7380

Qy   7381 GACTGAGCCGCGGGATGCGGGATGCGGGATCCGGGATCCGGGAGCTTAGAAAACAAAGTCACTATTATTAAATAGACAACCTCT 7440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7381 GACTGAGCCGCGGGATGCGGGATGCGGGATCCGGGATCCGGGAGCTTAGAAAACAAAGTCACTATTATTAAATAGACAACCTCT 7440

Qy   7441 GTTCCCCCATCATACCACAAGCAAGTCAGTTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7441 GTTCCCCCATCATACCACAAGCAAGTCAGTTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500

Qy   7501 CGGTGTCATGTGGGACTATGATGAAGTAGCAGTCAGCCCCAGGTCTCCTAAGTCTGTCAAGATC 7560
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7501 CGGTGTCATGTGGGACTATGATGAAGTAGCAGTCAGCCCCAGGTCTCCTAAGTCTGTCAAGATC 7560

Qy   7561 CCACATCACTGCCTTCGTTCGGGGCACTGAGTTCGTTCGTCTGGAGCCCGCAAGGCTGTTCT 7620
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7561 CCACATCACTGCCTTCGTTCGGGGCACTGAGTTCGTTCGTCTGGAGCCCGCAAGGCTGTTCT 7620

Qy   7621 GGACTTGCAGAAGTGTGTGCGAGGCAGGTTGTCTGGAGGTGAGATACCGAGTCATTAATCGGCAAACTGTGAT 7680
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7621 GGACTTGCAGAAGTGTGTGCGAGGCAGGTTGTCTGGAGGTGAGATACCGAGTCATTAATCGGCAAACTGTGAT 7680

Qy   7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAAACAACAAAGAAACCCCAAG 7740
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAAACAACAAAGAAACCCCAAG 7740

Qy   7741 GCTTATCTCGTACCCCACCTTGAAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGT 7800
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7741 GCTTATCTCGTACCCCACCTTGAAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGT 7800

Qy   7801 TGCTCCTGACGTAGTAGTAAAGCTGTCATGGGATGTGCTACGGGTTTGTAGATCCACGTAC 7860
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7801 TGCTCCTGACGTAGTAGTAAAGCTGTCATGGGATGTGCTACGGGTTTGTCGACCCACGTAC 7860

Qy   7861 CCGTGTCAAGGTCTCGTGTGCATGTGTCACCCGATGCAGTCGATGCCACATGCGATAC 7920
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7861 CCGTGTCAAGGTCTCGTGTGCATGTGTCACCCGATGCAGTCGATGCCACATGCGATAC 7920

Qy   7921 AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTGGAGACAGACATCTACTC 7980
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7921 AGTGTGTTTGACAGTACCATCACACCACGAGGATATCATGGTGGAGACAGACACATCTACTC 7980

Qy   7981 AGCAGCTAAACTCAGTGACCAACTCGGCATTCACACCATTGGCAGCAGTTATA 8040
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7981 AGCAGCTAAACTCAGTGACCAACTCGGCATTCACACCATTGGCAGCAGTTATA 8040

Qy   8041 CGCTGGAGGACCGATGCTTATGATGCCGAGAGATCGGATATCGGAGTCTAGGTC 8100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8041 CGCTGGAGGACCGATGCTTATGATGCCGAGATCGGATATCGGAGTCTAGGTC 8100

Qy   8101 TTCCGGGGTCTTATACTACCTCAAGTTTGACCTGCTGCTGGCTGAAGGTAAATGC 8160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8101 TTTCGGGGTCTTATACTACCTCAAGTTTGACCTGCTGCTGGCTGAAGGTAAATGC 8160

Qy   8161 TGCAGCCGAACAGGCTGGCATGCGGCCACCACAGATTCCCGCTTCCGTTATTTGCGCAC 8220
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8161 TGCAGCCGAACAGGCTGGCATGCGGCCACCACAGATTCCCGCTTCCGTTATTTGCGCAC 8220

Qy   8221 CGTAATTTGGAAGCAGCGCCGGAGATGCAGACAATGCAAGACAAAGCGTGTCTTTGCTAG 8280
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8221 CGTAATTTGGAAGCAGCGCCGGAGATGCAGACAATGCAAGACAAAGCGTGTCTTTGCTAG 8280

Qy   8281 CTGGATGAAGGTGATGGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8281 CTGGATGAAGGTGATGGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340

Qy   8341 AGAATAACATCATGCTCATCAAAGTGTTACCTCGAATTACCAAAAGTGGCCAAGCCTTA 8400
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8341 AGAATAACATCATGCTCATCAAAGTGTTACCTCGAATTACCAAAAGTGGCCAAGCCTTA 8400

Qy   8401 CTACTTTCTTACAAGAGATCTCGTATCCCCCTTGGGCGTGCTCGAGGGTCTGGG 8460
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8401 CTACTTTCTTACAAGAGATCTCGTATCCCCCTTGGGCGTGCTCGAGGGTCTGGG 8460

Qy   8461 ATACAACCCCAGTGTGCTGCGCGTGGATTGGGTATCTAATACATCACTACCCCCATGTTTGTGGGT 8520
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8461 ATACAACCCCAGTGTGCTGCGCGTGGATTGGGTATCTAATACATCACTACCCCCATGTTTGTGGGT 8520

```
Qy      8521 TAGCCGTGTGTTGGCTGCTGTCCATTTCATGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8521 TAGCCGTGTGTTGGCTGCTGTCCATTTCATGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580

Qy      8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640

Qy      8641 CATCATTGCTGGTGTGCACGGTATTGAGGCTTTCTCCGGTGGTGCGCTACCACCAACGCTGA 8700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8641 CATCATTGCTGGTGTGCACGGTATTGAGGCTTTCTCCGGTGGTGCGCTACCACCAACGCTGA 8700

Qy      8701 GATCCTCAGAGTTTCCAATCACTAACAGACATGACCATGCCCCCCTGCGAGCCTGGCG 8760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8701 GATCCTCAGAGTTTCCAATCACTAACAGACATGACCATGCCCCCCTGCGAGCCTGGCG 8760

Qy      8761 AAAGAAAGCCAGGGCGGTCCTCGCCAGCGCACGAGGCGTGGGAGCACACGCAAAATT 8820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8761 AAAGAAAGCCAGGGCGGTCCTCGCCAGCGCACGAGGCGTGGGAGCACACGCAAAATT 8820

Qy      8821 GGCTTCGCTTCCTTCTCTGGCATGCTACATCTAGACCTTCACCAGATTTGGATAAGACGAG 8880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8821 GGCTTCGCTTCCTTCTCTGGCATGCTACATCTAGACCTTCACCAGATTTGGATAAGACGAG 8880

Qy      8881 CGTGGCTTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGGAGGGGATGTGTT 8940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8881 CGTGGCTTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGGAGGGGATGTGTT 8940

Qy      8941 TATTACCACAGAGAAGATTGCAGAAGTTCCTTGTGAAGTATTGGCTGTCATTGTTTT 9000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8941 TGTTACCACAGAGAAGATTGCAGAAGTTCTTGTGAAGTATTGGCTGTCATTGTTTT 9000

Qy      9001 TGCCCTAGGGCTTCATTGCTGTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA 9060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9001 TGCCCTAGGGCTTCATTGCTGTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA 9060

Qy      9061 CTAACAG----TTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC 9116
             ||||||     ||||||||||||||||||||||||||||||||||||||||||||||||
Db      9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC 9120

Qy      9117 GGGCTTAACGACCCCGC 9133
             ||||||||||||||||
Db      9121 GGGCTTAACGACCCCGC 9137
```

RESULT 8

AR230431

LOCUS   AR230431   9143 bp   DNA   linear   PAT 20-DEC-2002
DEFINITION Sequence 390 from patent US 6451578.
ACCESSION   AR230431
VERSION   AR230431.1 GI:27270570
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 9143)
  AUTHORS   Simons,J.N., Pilot-Matias,T.J., Dawson,G.J., Schlauder,G.G.,
            Desai,S.M., Leary,T.P., Muerhoff,A.S., Erker,J.C., Buijk,S.L. and
            Mushahwar,I.K.
  TITLE   Non-A, non-B, non-C, non-D, non-E hepatitis reagents and methods
            for their use
  JOURNAL   Patent: US 6451578-A 390 17-SEP-2002;
FEATURES   Location/Qualifiers
     source   1..9143
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match   96.4%; Score 9059.8; DB 6; Length 9143;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 9096; Conservative   0; Mismatches   37; Indels   4; Gaps   1;

```
Qy      1 ACCACAAACACTCCAGTTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCCTAG 60
```

```
Db  1081  TTGGACTGGCACGGACTCCTTCTTGGCTGACCACATTGATTTGTTTATGGCGCTCTTGT  1140

Qy  1141  GACCTGTGACACGCCCTTGACATTGGTGAGTTGTGTGGCGTGTATTAGTCGGTTGACTG  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  GACCTGTGACACGCCCTTGACATTGGTGAGTTGTGTGGCGTGTATTAGTCGGTTGACTG  1200

Qy  1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTGTACTTGTTACCTGGA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTGTACTTGTTACCTGGA  1260

Qy  1261  AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGTGGATGGCCGGCAAGGT  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGTGGATGGCCGGCAAGGT  1320

Qy  1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT  1380

Qy  1381  GTTTAGCAGTGTACACTACCTGGCGGTTGCGCGCTCGATCTACTACACCTCTCCGGGGCAA  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  GTTTAGCAGTGTACACTACCTGGCGGTTGCGCGCTCGATCTACTACACCTCTCCGGGGCAA  1440

Qy  1441  GTGTATCAGTTGCTTCTCCTAGGCGCTTATGCTTTACATAGAAGCGACCTCTGAAAACCCCAT  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  GTGTATCAGTTGCTTCTCCTAGGCGCTTATGCTTTACATAGAAGCGACCTCTGAAAACCCTAT  1500

Qy  1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCTTTGAGTGATACCATGTCC  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCTTTGAGTGATACCATGTCC  1560

Qy  1561  TTGCCACTCTATTTGAGTGAGAAGTGTCTCAGAAGTCATTGTTACAGTCCAAAGTGGAC  1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  TTGCCACTCTATTTGAGTGAGAAGTGTCTCAGAAGTCATTGTTACAGTCCAAAGTGGAC  1620

Qy  1621  CAGGCCCTATCACTCTAGAGTATAACAACTCCATATCTTGTGCTACCCCTATACAATCCCTGG  1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  CAGGCCCTGTCACTCTAGAGTATAACAACTCCATATCTTGTGCTACCCCTATACAATCCCTGG  1680

Qy  1681  TGCGAGGGGATGTATGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  TGCGAGGGGATGTATGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA  1740

Qy  1741  TGTGCCATCGTACTGCGACTATGGGCACTGATGCAGTGTGAACGACACTCGCAACACTTA  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  TGTGCCATCGTACTGCGACTATGGGCACTGATGCAGTGTGAACGACACTCGCAACACTTA  1800

Qy  1801  CGAAGCATGCCGGTGTAACAACCATGGCGTAACAACGGACAACGGCTCAGCCCTGAA  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  CGAAGCATGCCGGTGTAACAACCATGGCGTAACAACGGACAACGGCTCAGCCCTGAA  1860

Qy  1861  ATTGGGCTATATTACAATACCCTGGGTCTGGAAAAATGTTTAAACCTCATATATTGGATGTC  1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  ATTGGGCTATATTACAATACCCTGGGTCTGGAAAAATGTTTAAACCTCATAATTGGATGTC  1920

Qy  1921  AGGCCATTTGTATTTGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAA  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  AGGCCATTTGTATTTGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAA  1980

Qy  1981  TTTCCACTCTCCTACCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCACCTGTGTACG  2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  TTTCCACTCTCCTACCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCACCTGTGTACG  2040

Qy  2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTACCACAGG  2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTACCACAGG  2100

Qy  2101  ATTGATCACCAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCATATTCGCGTAT  2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  ATTGATCACCAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCATATTCGCGTAT  2160

Qy  2161  TTTGTCTTTACGGGAGTTACCACCAAGGGCCGGTGGGTCTAATTCTGTTGGGTGTGTGG  2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TTTGTCTTTACGGGAGTTACCACCAAGGGCCGGTGGGTCTAATTCTGTTGGGTGTGTGG  2220

Qy  2221  CAGCAAGTATCTTATTTAGCCTACCTCTGTGTACTTCGTGTTTTGTTTGGCGGCGGCTTC  2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  CAGCAAGTATCTTATTTAGCCTACCTCTGTGTCCTCCGTGTTTTGTTTGGCGGCGGCTTC  2280

Qy  2281  TGGTTACCCTTTGCGTCCTGTCTCCCATCCCAGTCGTATCTCCAAGTGGCTGGGATGT  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TGGTTACCCTTTGCGTCCTGTGCTCCCGTATCTCCCAGTCGTATCTCCAAGTGGCTGGGATGT  2340

Qy  2341  TTTGTCTAAAGCTCAAGTCAAGTCCTTTGCTTGTTCTTCGATTTCTTCATCTGTGCTATCCTCCG  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  TTTGTCTAAAGCTCAAGTCAAGTCCTTTGCTTGTTCTTCGATTTCTTCATCTGTGCTATCCTCCG  2400

Qy  2401  CTGCAGGCTACGTTATCGTCGCCATGGTTGTGCCCATTGGGTTGTGCCCTGCGCGGCTGCCCCT  2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  CTGCAGGCTACGTTATCGTCGCCATGGTTGTGCCCATTGGGTTGTGCCCTGCGCGGCTGCCCCT  2460

Qy  2461  AACTTTCTTGTTCGAGCAGCTGCCTTTAACCCCTTGCATTGGTTACGCCTTCAGC  2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  AACTTTCTTGTTCGAGCAGCTGCCTTTAACCCCTTGCATTGGTTACGCCTTCAGC  2520

Qy  2521  AGTGGCAGGGTTAGTTTTGTGTGGGCCGGCCGTAACCCGTGTCACCCGTTGCTCTTGT  2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  AGTGGCAGGGTTAGTTTTGTGTGGGCCGGCCGTAACCCGTGTCACCCGTTGCTCTTGT  2580

Qy  2581  AGTCCTGGCCTCTGCGAGCAGCTTTTAACCCCTTGCATTGGCTACGCCTGCTTCAGC  2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  AGTCCTGGCCTCTGCGAGCAGCTTTTAACCCCTTGCATTGGCTACGCCTGCTTCAGC  2640

Qy  2641  TTTTGATACCGAGATAATTGGAGGGCTGACATATCACCCTGTAGTAGCATTAGTTGTCAT  2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  TTTTGACACCGAGATAATTGGAGGGCTGACATATCACCCTGTAGTAGCATTAGTTGTCAT  2700

Qy  2701  GTCTCGTTTGCTTCTTGCTCACTTGTACCCTGCGTGCTTACCTGCGTGTACTAAGACCGGAGGTTTTCCT  2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  GTCTCGTTTGCTTCTTGCTCACTTGTACCCTGCGTGCTTACCTGCGTGTACTAAGACCGGAGGTTTCTCCT  2760

Qy  2761  TTGGCAACGTTGGGAGAATTGGTTTCCCCGGTGCGACATATGACGCGGTGGTACTTTCTGTGTGTCA  2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  TTGGCAACGTTGGGAGAATTGGTTTCCCCGGTGCGACATATGACGCGGTGGTACTTTCTGTGTGTCA  2820

Qy  2821  TGTGCTGGTTGTTTTCCCCGGTGCACTGATGCAGTGACGACACGGTGGTACTTTCTGTGTGTCA  2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  TGTGCTGGTTGTTTTCCCCGGTGCACTGATGCAGTGACGACACGGTGGTACTTTCTGTGTGTCA  2880

Qy  2881  CGTAGCTCTTCTTCATGTTTAACATCCAGTGCGCTCGAGCATCGTCTTCTTGGGACTGACTCTAGGGT  2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  CGTAGCTCTTCTTCATGTTTAACATCCAGTGCGCTCGAGCATCGTCTTCTTGGGACTGACTCTAGGGT  2940

Qy  2941  TAGGGCCCATAGAATGTTGGTGCGTCTCAGAAAGTGTCATGCTTGGTGGTATTCTCATTATGT  3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941  CGTAGGGCCCATAGAATGTTGGTGCGTCTCAGAAAGTGTGAGAATGGTGTGTTCTCATTATGT  3000

Qy  3001  TCTTAAGTTTTCCTCTTCAGTGTTTGGTTGGTGAGAATGGTGTGTTTTTCTATAAGCACTTGCA  3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  TCTTAAGTTTTCCTCTTCAGTGTTTGGTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA  3060

Qy  3061  TGGTGATGTCTTGCTCAATGATTTTGCCTCGAAACTTGCAAGAGCCATTGTTCCC  3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061  TGGTGATGTCTTGCTCAATGATTTTGCCTCGAAACTTGCAAGAGCCATTGTTTCCC  3120

Qy  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAAGTGAAGGAAGACGCTTGGCGTGTGGGGACAC  3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAAGTGAAGGAAGACGCTTGGCGTGTGGGGACAC  3180

Qy  3181  GGTTGATGGTTTTGCCCGTCTTGTTTCGGCGACCTTGTTTCCGCAGGGTTGGCTAT  3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181  GGTTGATGGTTTTGCCCGTCTTGTTTCGGCGACCTTGTTTCCGCAGGGTTAGCTAT  3240

Qy  3241  GCCGCCAGATGGGTGGGCCATTACCCGTCTGGTCTAATTCTGTTGTTGGGGTTGTGTGG  3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241  GCCGCCAGATGGGTGGGCCATTACCCGTCTGGTCTAATTCTGTTGTTGGGGTTGTGTGG  3300
```

Right column:

```
Qy   4381 TGATGAGCTTGCTAACGAGTTAGCTGGAATAACAGCTGTCTCTTACTATAGGGG 4440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 TGATGAGCTTGCTAACGAGTTAGCTGGAATAACAGCTGTCTCTTACTATAGGGG 4440

Qy   4441 ATGTGACATCTCAAAAATCCCTGAGGGGCCACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 ATGTGACATCTCAAAAATCCCTGAGGGGCCACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500

Qy   4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTTGCAGCCTCATGGTAGAAGGCAC 4560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTTGCAGCCTCATGGTAGAAGGCAC 4560

Qy   4561 ATGGCCATGTGACCTTGACCCTACTTTCACCATGGTGTCGTGTGCGGGGTCTCAGC 4620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 ATGGCCATGTGACCTTGACCCTACTTTCACCATGGTGTCGTGTGCGGGGTCTCAGC 4620

Qy   4621 AATAGTTAAAGGCCAGCGTAGGGGCTGGGGAGAGCTGGCATATACTACTA 4680
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   4621 AATAGTTAAAGGCCAGCGTAGGGGCTGGGGAGAGCTGGCATATACTACTA 4680

Qy   4681 TGTAGACGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4681 TGTAGACGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740

Qy   4741 CGACGCAGCCAAGGCATGTGATGGTTGTCATCAACAGAGCTCAAACTATTCTGGACAC 4800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 CGACGCAGCCAAGGCATGTGATGGTTGTCATCAACAGAGCTCAAACTATTCTGGACAC 4800

Qy   4801 CTATCGCACCACCTGGGTTACCTGCGATAGGAGCAAATTGGACGAGTGGGCTGATCT 4860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 CTATCGCACCACCTGGGTTACCTGCGATAGGAGCAAATTGGACGAGTGGGCTGATCT 4860

Qy   4861 CTTTTCATGTGTCAACCCGAACCTTCATTGTCAATACTGCTGACAA 4920
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   4861 CTTTTCATGTGTCAACCCGAACCTTCATTGTCAATACTGCTGACAA 4920

Qy   4921 TTATGTTTTGTTGACTGCGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGCTCTCC 4980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 TTATGTTTTGTTGACTGCGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGCTCTCC 4980

Qy   4981 CAATGACGCACCACGGTGGCAGGGAGCCCGCTTGTGGGGAAAAACCTTGTGGGGTTCTGTG 5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 CAATGACGCACCACGGTGGCAGGGAGCCCGCTTGTGGGGAAAAACCTTGTGGGGTTCTGTG 5040

Qy   5041 GCGCTTGACGGCGGCCTGCGACGCGTCCTGGCCCAGAGCCCAGCGAGGTGACCAGATACCA 5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041 GCGCTTGACGGCGGCCTGCGACGCGTCCTGGCCCAGAGCCCAGCGAGGTGACCAGATACCA 5100

Qy   5101 AATGTGCTTCACTGAAGTCAATACTTCTGGGACAGCCCGCACTCGCTGTGTCGCTGCCACTGTGCTCC 5160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101 AATGTGCTTCACTGAAGTCAATACTTCTGGGACAGCCCGCACTCGCTGTGTCGCTGGAGT 5160

Qy   5161 GGCTATGGTTATCTAGCCATTGGCCAGTCCTTCATTGACACTTTTGGCGCGTTGCTGGTC 5220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161 GGCTATGGTTATCTAGCCATTGGCCAGTCCTTCATTGACACTTTTGGCGCGTTGCTGGTC 5220

Qy   5221 TATTACATCAGTCCCTACTGTCGCCCCAGTGTGTTGACGAGAAGAAATCGT 5280
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221 TATTACATCAGTCCCTACTGTCGCCCCAGTGTGTTGACGAGAAGAAATCGT 5280

Qy   5281 GGAGGAGTTGCATCATTCATTCCCTGAGTAATCCGATAAGCTGAA 5340
          |||||||||||||||||||||||||||||||||||||||||||||
Db   5281 GGAGGAGTTGCATCATTCATTCCCTGAGTAATCCGATAAGCTGAA 5340

Qy   5341 GAGTACAATCACCACAACTAGTCTTCACATTGGAAACCGCCCTGAAAAACTTAACAC 5400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5341 GAGTACAATAACCACAACTAGTCTTCACATTGGAAACCGCCCTGAAAAACTTAACAC 5400

Qy   5401 CTTTCTTGGGCCTCATGCAAGTCCTACAATCCTTGCTATCATAGATATTGCTGTGGTTTAGT 5460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5401 CTTTCTTGGGCCTCATGCAAGTCCTACAATCCTTGCTATCATAGATATTGCTGTGGTTAGT 5460

Qy   5461 CACTTACCTGACAATCCTTTGCATCATGTCGTGTTGCTTTCATTGCGGGTATTACTAC 5520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5461 CACTTACCTGACAATCCTTTGCATCATGTCGTGTTGCTTTCATTGCGGGTATTACTAC 5520
```

Left column:

```
Qy   3301 CACGCTGTCAGCCGATGGCAGTGGTCATGACTGGTATAGACCCCGACTTGGACTGGAAC 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 CACGCTGTCAGCCGATGGCAGTGGTCATGACTGGTATAGACCCCGACTTGGACTGGAAC 3360

Qy   3361 TATCTTCAGATTAGGATCTCTGCCACTAGCTACATGGGATTTGTTTGTGACAACGTGTT 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 TATCTTCAGATTAGGATCTCTGCCACTAGCTACATGGGATTTGTTTGTGACAACGTGTT 3420

Qy   3421 GTATACTGCTCACCATGGCAGCAGGCGGCCCGTTGGCTCATCCCACAGGCTCTATACA 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 GTATACTGCTCACCATGGCAGCAGGCGGCCCGTTGGCTCATCCCACAGGCTCCATACA 3480

Qy   3481 CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG 3540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG 3540

Qy   3541 GTCCCTTACTCGGTGCTCTTGCCGGTGGAGACCAAGGGTATCTGGTACACGACTGGGGTC 3600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 GTCCCTTACTCGGTGCTCTTGCCGGTGGAGACCAAGGGTATCTGGTACACGACTGGGGTC 3600

Qy   3601 AATTGGTTGAGGTGCAACAAATCCGATGACCCTTATTGGTGTGTGCGGGGCCCTTCCCAT 3660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 AATTGGTTGAGGTGCAACAAATCCGATGACCCTTATTGGTGTGTGCGGGGCCCTTCCCAT 3660

Qy   3661 GGCTGTTGCCAAGGGTTCTTGCAGTGCGCCCGATTCTGTGCTCCTCCGGGCATGTTATTGG 3720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 GGCTGTTGCCAAGGGTTCTTGCAGTGCGCCCGATTCTGTGCTCCTCCGGGCATGTTATTGG 3720

Qy   3721 GATGTTCACCGTGCTCTAGAAATTCTGGCGGTTCAGTCAGTATTAGGGTTAGGCCGTT 3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 GATGTTCACCGTGCTCTAGAAATTCTGGCGGTTCAGTCAGTATTAGGGTTAGGCCGTT 3780

Qy   3781 GGTGTGTGCTCGGATACCATCCCAGTACACAGCACATGCCCACTCCTTGATACAAACCTAC 3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 GGTGTGTGCTCGGATACCATCCCAGTACACAGCACATGCCCACTCCTTGATACAAAACCTAC 3840

Qy   3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCACGGCAAGTCAAC 3900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCACGGCAAGTCAAC 3900

Qy   3901 CAAATTACCACTTTCTTACATGGAGGAAGTATGAGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT 3960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 CAAATTACCACTTTCTTACATGGAGGAAGTATGAGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT 3960

Qy   3961 GGCTACAACAGCATCAATGCCAAAAGTACTGCCGAAAAGTACGCGACGTACGCGCCAAATTG 4020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 GGCTACAACAGCATCAATGCCAAAAGTACTGCCGAAAAGTACGCGACGTACGCGCCAAATTG 4020

Qy   4021 CTATTTTAATGGCAAATGTACCAACACAGGGGCCTCACTTACGTACAGCACATATGGCAT 4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 CTATTTTAATGGCAAATGTACCAACACAGGGGCCTCACTTACGTACAGCACATATGGCAT 4080

Qy   4081 GTACCTGACCGGAGCATGTTCCCGGACTATGACGTAATCATTTGTGACCGAATGCCATGC 4140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 GTACCTGACCGGAGCATGTTCCCGGACTATGACGTAATCATTTGTGACCGAATGCCATGC 4140

Qy   4141 TACCGATGCAACCACCGTGTTGGGCATTGGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA 4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 TACCGATGCAACCACCGTGTTGGGCATTGGGAAAGGTTCTAACCGAAGCTCCATCCAAAAA 4200

Qy   4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCTGAGTAATCCTGAGCACACATGC 4260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCTGAGTAATCCTGAGCACACATGC 4260

Qy   4261 CAACATAACTGAGATTCAATTAACCGTAATAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAGAT 4320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 CAACATAACTGAGATTCAATTAACCGTAATAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAGAT 4320

Qy   4321 TAAGGAGAAAAATCTGAAGAAAGGGAGACACCTTATCTTTGAGGGTACCAAAAAACACTG 4380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321 TAAGGAGAAAAATCTGAAGAAAGGGAGACACCTTATCTTTGAGGGTACCAAAAAACACTG 4380

Qy   4381 TGATGAGCTTGCTAACGAGTTAGCTGGAATAACAGCTGTCTCTTACTATAGGGG 4440
```

Db

```
Qy  5461 CACTTTACCTGACAATCCCTTTGCATCATGCGTGTTTGCTTTCATTGCGGTATTACTAC 5520
Db  5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTTGGAGGGCAATTGCGTCCAA 5580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTTGGAGGGCAATTGCGTCCAA 5580

Db  5581 GCTTACAGACGCTAGAGGCCGCACTGGCGTTCATGATGGCCGGGGCGTGCGCGGAACAGCTCT 5640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  5581 GCTTACAGACGCTAGAGGCCGCACTGGCGTTCATGATGGCCGGGGCGTGCGCGGAACAGCTCT 5640

Db  5641 TGGTACATGGACATCGGTGGGTTTGCTTGTCTTTGACATGCTCTAGGCGGCGTATGCTGCCGCTC 5700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  5641 TGGTACATGGACATCGGTGGGTTTGCTTGTCTTTGACATGCTCTAGGCGGCGTATGCTGCCGCTC 5700

Db  5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGGCCCACTATGGATCAGCT 5760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGGCCCACTATGGATCAGCT 5760

Db  5761 TGCTGGTTTAGTCTACTCCGCGTTCAATCCGGCCCCGAGGAGTTGTGGCGTCTTGTCAGC 5820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  5761 TGCTGGTTTAGTCTACTCCGCGTTCAATCCGGCCCCGAGGAGTTGTGGCGTCTTGTCAGC 5820

Db  5821 TTGTGCAAATGTTTGCTTGACAACAGCAGGGCCAGAATCACTGGCCCAACAGACTTCTTAC 5880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  5821 TTGTGCAAATGTTTGCTTGACAACAGCAGGGCCAGAATCACTGGCCCAACAGACTTCTTAC 5880

Db  5881 TATGCTTGCTTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  5881 TATGCTTGCTTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940

Db  5941 CAGGAAGATACTGGGCATTCTGGAGCATCTACCCCCTGAGTGTCATATCAGCGTTGCAT 6000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  5941 CAGGAAGATACTGGGCATTCTGGAGCATCTACCCCCTGAGTGTCATATCAGCGTTGCAT 6000

Db  6001 CCGTTGGCTCCACACCCCCACACCCCGCGAGGATGATTGCGGCTCATTGCTTGGGGTCTAGAGAT 6060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6001 CCGTTGGCTCCACACCCCCACACCCCGCGAGGATGATTGCGGCTCATTGCTTGGGGTCTAGAGAT 6060

Db  6061 TTGGCAGTATGTGTGCAATTTCTTTGTGATTGCTTTAATGTCCTTAAAGCTGGAGTTCA 6120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6061 TTGGCAGTATGTGTGCAATTTCTTTGTGATTGCTTTAATGTCCTTAAAGCTGGAGTTCA 6120

Db  6121 GAGCATTCAGGTATGCCTCCAAGCACGCTGTCCAGAAGACGGGGTACAAGGGCCC 6180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6121 GAGCATTCAGGTATGCCTCCAAGCACGCTGTCCAGAAGACGGGGTACAAGGGCCC 6180

Db  6181 CTTGAGAGCTTACAAAGACCCCAGAACTTGTTCAAATTACTGGAG 6240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6181 CTTGAGAGCTTACAAAGACCCCAGAACTTGTTCAAATTACTGGAG 6240

Db  6241 TGTTGAGAAATGTTTTGCAAAACTTTACACAGGACCCAACTTGTTCAAATTACTGGAG 6300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6241 TGTTGAGAAATGTTTTGCAAAACTTTACACAGGACCCAACTTGTTCAAATTACTGGAG 6300

Db  6301 AGGGGCTGTTCCAGTCAACGCGTAGGCGTGTGTGGGTCGGCTAGACCCCAACTGATTG 6360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6301 AGGGGCTGTTCCAGTCAACGCGTAGGCGTGTGTGGGTCGGCTAGACCCCAACTGATTG 6360

Db  6361 GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGGTAAATATGAGAAAATGGGGAGA 6420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6361 GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGGTAAATATGAGAAAATGGGGAGA 6420

Db  6421 TCACATTTTTGTTACAGCAGTATCCTCTCCAAATGTCGTTTCACCCAGGTGCCCCCAAC 6480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6421 TCACATTTTTGTTACAGCAGTATCCTCTCCAAATGTCGTTTCACCCAGGTGCCCCCAAC 6480

Db  6481 CTTGAGAGCTGCAGTGCCGTGACGCGTACAGGTTCAGTGTTATCTGAGCCCAA 6540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6481 CTTGAGAGCTGCAGTGCCGTGACGCGTACAGGTTCAGTGTTATCTGAGCCCAA 6540

Db  6541 AACTCCCTTGGACGACATCTGCTTGCTTGCTGTTACGGTCCGACGGTAAGGTAAAACTGTTAA 6600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6541 AACTCCCTTGGACGACATCTGCTTGCTTGCTGTTACGGTCCGACGGTAAGGTAAAACTGTTAA 6600
```

Db

```
Qy  6601 GCTTCCCTTCCGCGTTGACGGTCACACCTTGTGTGCCGCATGCAACTTAATTGCGTGA 6660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601 GCTTCCCTTCCGCGTTGACGGTCACACCTTGTGTGCCGCATGCAACTTAATTGCGTGA 6660

Qy  6661 TGCACTTGAGACAAATGACTGTAATTCCACAACAACACTCCAGTGATGAGCCGCAGT 6720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661 TGCACTTGAGACAAATGACTGTAATTCCATAAACAACAACTCCAGTGATGAGCCGCAGT 6720

Qy  6721 GTCCGCTCTTGTTTTCAAACAGGAGTTGCGCGTACAAACCAATTGCTTGAGGCAATTTC 6780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721 GTCCGCTCTTGTTTTCAAACAGGAGTTGCGCGGCTACAAACCAATTGCTTGAGGCAATTTC 6780

Qy  6781 AGCTGGCGTTGACACCACCAAACTGCCAGCCCCCTCATCGAACGCCCCCTCCGAGATCCGTCCC 6840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781 AGCTGGCGTTGACACCACCAAACTGCCAGCCCCCTCATCGAACGAGGTAGTGTAAGAAA 6840

Qy  6841 GGGCCAGTTCCGGCAAGAACTGGTTCGCTTACACTCCTTGCCTCCCCTGCGTCCCCCCTCC 6900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841 GGGCCAGTTCCGGCAAGAACTGGTTCGCTTACACTCCTTGCCTCCCCTGCGTCCCCCCTCC 6900

Qy  6901 AGGAGTGTCATGTCCTGAAAGCTGCAACGAAGTGACCCGTTAGAAGTGTCCTTCAAACCT 6960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901 AGGAGTGTCATGTCCTGAAAGCTGCAACGAAGTGACCCGTTAGAAGTGTCCTTCAAACCT 6960

Qy  6961 CCCTTCTTCACCACTGTTCTACAGTTGGCATGGCCATGCCGATGCCGTGTTGGGGACGGTGA 7020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961 CCCTTCTTCACCACTGTTCTACAGTTGGCATGGCCATGCCGATGCCGTGTTGGGGCAGGTGA 7020

Qy  7021 GTGTAACCCTTTCACTGCAATTGGATGGCGCAATGACCGGAAACAGGCCGGAGGCCCTGATGA 7080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGGAAACAGGCCGGAGGCCCTGATGA 7080

Qy  7081 TTTACCCAGTTACCCTCCCAAAAAGGAGGTCTCTGAATGGTCAGACGAAAGTTGGTCGAC 7140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081 TTTACCCAGTTACCCTCCCAAAAAGGAGGTCTCTGAATGGTCAGACGAAGTTGGTCAAC 7140

Qy  7141 GGCTACAACCGGCTTCCAGCTACTGGCCCCCCGTACCTGGCCTTACCTAAGATACGGGGAAAGGA 7200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141 GACTACAACCGGCTTCCAGCTACTGGCCCCCCGTACCTGGCCTTACCTAAGATACGGGGAAAGGA 7200

Qy  7201 TTCCACTCAGTCAGCCCCCCGCCAAACGGCCCTACAACGGTGGTTTCCTCAAAGAAGTTGGGAAGAGTGAGTT 7260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201 TTCCACTCAGTCAGCCCCCCGCCAAACGGCCCTACAACGGTGGTTTCCTCAAAGAAGTTGGGAAGAGTGAGTT 7260

Qy  7261 TTCGTGCAACTCGGCCATGAGCTACCTGGACCGATTAGCTTCAAAACTGCTTCTCTAAAGT 7320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261 TTCGTGCAACTCGGCCATGAGCTACACTGGACCGACGTGATTAGCTTCAAAACTGCTTCTCTAAAGT 7320

Qy  7321 TCTGTCTGCAACTCGGGCGATGCGGGATGCGGGGGTCTGTCAACTAGTGGTTTCCTCAAACAAGATCATTGGTATGTATGT 7380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7321 TCTGTCTGCAACTCGGGCGATGCGGGATGCGGGGGTCTGTCAACTAGTGGTTTCCTCAAACAAGATCATTGGTATGT 7380

Qy  7381 GACTGAGCGGCGGGATGCGGGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCT 7440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381 GACTGAGCGGCGGGATGCGGGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCT 7440

Qy  7441 GTTCCCCCCATCATCACCACAACGAGATTGGCTAAGGAAAAATGGGGAGA 7500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441 GTTCCCCCCATCATCACCACAACGAGATTGGCTAAGGAAAAATGGGGAGA 7500

Qy  7501 CGGTGTCATGTGGGACTATGAAGTGACTGAGTAGCAGCTCACACGCTCCTCTACGTCC 7560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7501 CGGTGTCATGTGGGACTATGAAGTGAGTAGCAGCTCACACGCTCCTCTACGTCTCTAAGTC 7560

Qy  7561 CCACATCACTGGCCTTCGGGGACTGATGTTCGTTGGAGCAGCCCCAAGGCTGTTCT 7620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7561 CCACATCACTGGCCTTCGGGGACTGATGTTCGTTCTGGAGCAGCCGCCAAGGCTGTTCT 7620

Qy  7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATAACCGGAGTCATTATCGGCAAACTGTGAT 7680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7621 GGACTTGCGAGAAGTGTGTCGAGGCAGGTGAGATAACCGGAGTCATTATCGGCAAACTGTGAT 7680
```

```
Qy   7681 AGTTCCAAAGGAGGAGAGTCTTCGTGAAGACCCCCAGAAACCAACAACAAAGAAACCCCCAAG 7740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7681 AGTTCCAAAGGAGGAGAGTCTTCGTGAAGACCCCCAGAAACCAACAACAAAGAAACCCCCAG 7740

Qy   7741 GCTTATCTCGTACCCCACCTTGAAATGATGTTGAGAGATGTACTACGGTCAGGT 7800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7741 GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAGATGTACTACGGTCAGGT 7800

Qy   7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTACGGGTTGTTGTAGATCCACGTAC 7860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTACGGGTTGTTGTAGATCCACGTAC 7860

Qy   7861 CCGTGTCAAGCGTCTGTGTCGATGTGGTCACCCGGATCGGAGCCACATGCGATAC 7920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7861 CCGTGTCAAGCGTCTGTGTCGATGTGGTCACCCGGATCGGAGCCACATGCGATAC 7920

Qy   7921 AGTGTGTTTGACAGTACCATCACACACCCGAGGATATCATGGTGGAGACAGACATCTACTC 7980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7921 AGTGTGTTTGACAGTACCATCACACACCCGAGGATATCATGGTGGAGACAGACATCTACTC 7980

Qy   7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCATTCACACCATTGCGAGGCAGTATA 8040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCATTCACACCATTGCGAGGCAGTATA 8040

Qy   8041 CGCTGGAGGACCGATGATCGCTTATGATGGCCCAGAGATCGGATATCCTAGGTGTAGGTC 8100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8041 CGCTGGAGGACCGATGATCGCTTATGATGGCCCAGAGATCGGATATCCTAGGTGTAGGTC 8100

Qy   8101 TTCCGGGGTCTATACTACCTCAAGTTCAACAGTTGACCTGCTGGCTGAAGTAAATGC 8160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8101 TTCCGGGGTCTATACTACCTCAAGTTCAACAGTTGACCTGCTGGCTGAAGTAAATGC 8160

Qy   8161 TGCAGCCGAACAGGCTGCATGAAGAGAACCCTGCGCTTCGCTTATTTGCGGCGATGATTGCAC 8220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8161 TGCAGCCGAACAGGCTGCATGAAGAGAACCCTGCGCTTCGCTTATTTGCGGCGATGATTGCAC 8220

Qy   8221 CGTAATTGGAAGAGCGCCGGAGAGCGCCGGAGATGCAGACAAACAAGCAATGCGTGTCTTTGCTAG 8280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8221 CGTAATTGGAAGAGCGCCGGAGAGCGCCGGAGATGCAGACAAACAAGCAATGCGTGTCTTTGCTAG 8280

Qy   8281 CTGGATGAAGGTGATGGCGTGACCACCAGATTGTGTGCCTCAAACCCAAATACAGTTTGGA 8340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8281 CTGGATGAAGGTGATGGCGTGACCACCAGATTGTGTGCCTCAAACCCAAATACAGTTTCGA 8340

Qy   8341 AGAATTAACACATGCTCATCAAAATGTACCTCGGAATTACCAAAAGTAGTGGCAAGCCTTA 8400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8341 AGAATTAACACATGCTCATCAAAATGTACCTCGGAATTACCAAAAGTAGTGGCAAGCCTTA 8400

Qy   8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTGCCGAGGGTCTGGG 8460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTGCCGAGGGTCTGGG 8460

Qy   8461 ATACAACCCAGTGCTGCGCGTGGCGTGGGATTCTAATCACTCTAATACACTACCCATGTTTGTGGGT 8520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8461 ATACAACCCAGTGCTGCGCGTGGCGTGGGATTCTAATCACTCTAATACACTACCCATGTTTGTGGGT 8520

Qy   8521 TAGCCGTGTGTTGGCGTGTCCATTCAGTGGAGCAGAGATGCTCTTTGAGGACAAAACTTCCCGA 8580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8521 TAGCCGTGTGTTGGCGTGTCCATTTCAGTGGAGCAGAGATGCTCTTTGAGGACAAAACTTCCCGA 8580

Qy   8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640

Qy   8641 CATCATTGCTCGGTGTGCACGGTATTGAGGCTTTCTCGGTGTGCGCTACCACCAGCGTGA 8700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8641 CATCATTGCTCGGTGTGCACGGTATTGAGGCTTTCTCGGTGTGCGCTACCACCAGCGTGA 8700

Qy   8701 GATCCTCCAGAGTTCCCAATCACTACTAACACTAACACTACCAACATGACCCCGTGCGGAGCCTGGCG 8760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8701 GATCCTCCAGAGTTCCCAATCACTACTAACACTAACACTACCAACATGACCCCGTGCGGAGCCTGGCG 8760

Qy   8761 AAAGAAAGCCAGGGCGGTCCTCGCACGTGGCCGGAGCACACGCCGCAAAATT 8820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8761 AAAGAAAGCCAGGGCGGTCCTCGCACGTGGCCGGAGCACACGCCGCAAAATT 8820

Qy   8821 GGCTCGCTTCCTTCTCTGGCATGCGACTCTAGACCTCTACACGATTTGGATAAGAGACGAG 8880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8821 GGCTCGCTTCCTTCTCTGGCATGCGACTCTAGACCTCTACACGATTTGGATAAGAGACGAG 8880

Qy   8881 CGTGGCTCGGTACACCACTTCAATTATTGTGATATTGTGATGTTACTCCCGGAGGGGATGTGTT 8940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8881 CGTGGCTCGGTACACCACTTCAATTATTGTGATATTGTGATGTTACTCCCGGAGGGGATGTGTT 8940

Qy   8941 TATTACACCACAGAGAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTT 9000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8941 TGTTACACCACAGAGAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTT 9000

Qy   9001 TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATTCAGCTGAACCCCCAAATTCAAAATTAA 9060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9001 TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATTCAGCTGAACCCCCAAATTCAAAATTAA 9060

Qy   9061 CTAACAG----TTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGCAACAGGGGAGACCCC 9116
          ||||||    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGCAACAGGGGAGACCCC 9120

Qy   9117 GGGCTTAACGACCCCGC 9133
          |||||||||||||||||
Db   9121 GGGCTTAACGACCCCGC 9137
```

RESULT 9
AR230434

```
LOCUS       AR230434                      9143 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 393 from patent US 6451578.
ACCESSION   AR230434
VERSION     AR230434.1  GI:27270573
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 9143)
  AUTHORS   Simons,J.N., Pilot-Matias,T.J., Dawson,G.J., Schlauder,G.G.,
            Desai,S.M., Leary,T.P., Muerhoff,A.S., Erker,J.C., Buijk,S.L. and
            Mushahwar,I.K.
  TITLE     Non-A, non-B, non-C, non-D, non-E hepatitis reagents and methods
            for their use
  JOURNAL   Patent: US 6451578-A 393 17-SEP-2002;
FEATURES             Location/Qualifiers
     source          1..9143
                     /organism="unknown"
                     /mol_type="genomic DNA"

ORIGIN
Query Match       96.4%;  Score 9059.8;  DB 6;  Length 9143;
Best Local Similarity   99.6%;
Matches 9096;  Conservative    0;  Mismatches   37;  Indels   4;  Gaps   1;
```

```
Qy      1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCTCTGGAGCACCCCCCTAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG 60

Qy     61 CAGGGCGTGGGGATTTCCCCGCCCGTCGCGAGAAGGTGGAGCCAACCACCTTAGTAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAGGGCGTGGGGATTTCCCCGCCCGTCGCGAGAAGGTGGAGCCAACCACCTTAGTAT 120

Qy    121 GTAGGCGGCGGACTCATGACGCTCGCTGTGATGACAAGCGCCAAGCTTGACTTGGATGGC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTAGGCGGCGGACTCATGACGCTCGCTGTGATGACAAGCGCCAAGCTTGACTTGGATGGC 180

Qy    181 CCTGATGGGCTTCATGGCGTTCATGGGTGGTGCAGCCTCCACGCCGCCACCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCTGATGGGCTTCATGGCGTTCATGGGTGGTGCAGCCTCCACGCGCCCACCA 240

Qy    241 CCTTCCCAGATAGAGCGCGGCGGCGCACTGTAGGGGAAGAACCGGGGAGACTACCAAGACG 300
```

```
Db  241 CCTCCCAGATAGAGCGGCGCACTCTGAGGAGGAAGACCGGGACCGGTCACTACCAAGGACG 300

Qy  301 CAGACCTCTTTTGAGTATCACGCCTTCCGGAAGTAGTTGGGCAAGCCACCTATATGTGT 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGACCTCTTTTGAGTATCACGCCTTCCGGAAGTAGTTGGGCAAGCCACCTATATGTGT 360

Qy  361 TGGGATGGTTGGGGTTGGGGTTAGCCATCCATACCGTACTACCGTACTACCGTACTACCG 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 TGGGATGGTTGGGGTTGGGGTTAGCCATCCATACCGTACTACCGTACTACCGTACTACCG 420

Qy  421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTGTTATTTCTACTCAAACAAGTCCTGTACC 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTGTTATTTCTACTCAAACAAGTCCTGTACC 480

Qy  481 TGCGCCCAGAACGCGCAAGAACAAGCAGACGACGAGGCTTCATATCCTGTGTCCATTAAAAC 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TGCGCCCAGAACGCGCAAGAACAAGCAGACGACGAGGCTTCATATCCTGTGTCCATTAAAAC 540

Qy  541 ATCTGTTGAAAGGGGACAACGACAAAGCCGCAAAGTCCAGCGCGATGCTCGGCCTCGTAA 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 ATCTGTTGAAAGGGGACAACGACAAAGCCGCAAAGTCCAGCGCGATGCTCGGCCTCGTAA 600

Qy  601 TTACAAAATTGCTCGTGATCGTGATGATGGCTTGCAGACATTGGCTCAGGCTGCTTGCCAGC 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 TTACAAAATTGCTCGTGATCGTGATGATGGCTTGCAGACATTGGCTCAGGCTGCTTGCCAGC 660

Qy  661 TCATGGTTGGGGACGGCGCAAGACCCTGCCAATCTTCGCAATCTTTGGAATCCTTCTGGA 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TCATGGTTGGGGACGGCGCAAGACCCTGCCAATCTTCGCAATCTTTGGAATCCTTCTGGA 720

Qy  721 TTACCCTTTGGGGTGGTGATTGGTGATGTTACAACTCACACACCTCTAGTAGGCCCGCTGGT 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 TTACCCTTTGGGGTGGTGATTGGTGATGTTACAACTCACACACCTCTAGTAGGCCCGCTGGT 780

Qy  781 GGCAGGAGCGGCGTCGTTCGACCAGTCTGCAGATAGTACGCTTGCTGGAGGATGGAGTCAA 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GGCAGGAGCGGCGTCGTTCGACCAGTCTGCAGATAGTACGCTTGCTGGAGGATGGAGTCAA 840

Qy  841 CTGTGATCTGTTGGTTCGCGGTTCCACCTTTTGTGTCTATCTTTGGCCTG 900
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CTGGGATCTGTTGGTTCGCGGTTCCACCTTTTGTGTCTATCTTTGGCCTG 900

Qy  901 TCCCTGTAGTGGGGCGCGGGTCACTGACCGGGTCACTGACCACAATACCACAAATTG 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TCCCTGTAGTGGGGCGCGGGTCACTGACCGGGTCACTGACCACAATACCACAAATTG 960

Qy  961 CTGCCAGCGTAATCAGGTTATTATCTGTTCCTCCTTCCACTTGCCTACACGAGCCTGTTG 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 CTGCCAGCGTAATCAGGTTATTATCTGTTCCTCCTTCCACTTGCCTACACGAGCCTGTTG 1020

Qy 1021 TGTGATCTGTGCGGACGAGTGCTGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 TGTGATCTGTGCGGACGAGTGCTGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA 1080

Qy 1081 TTGGACTGGCACGGACTCCTTCTTGGCTGACCGACCACACAGCTCAATTGGTGTACTG 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 TTGGACTGGCACGGACTCCTTCTTGGCTGACCGACCACACAGCTCAATTGGTGTGACTG 1140

Qy 1141 GACCTGTGACGGCCCTTGACATTGGTGAGTTGGTGGCGTGTTGGTGCGTGTTGGTGTGACTG 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 GACCTGTGACGGCCCTTGACATTGGTGAGTTGGTGGCGTGTTGGTGCGTGTTGGTGTGACTG 1200

Qy 1201 GCTTGTCAGGCAGGCACTGGCTTATTCAATGAAACTGGTACTTGTTGTTACCTGGA 1260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 GCTTGTCAGGCAGGCACTGGCTTATTCAATGAAACTGGTACTTGTTGTTACCTGGA 1260

Qy 1261 AGTGCCCACTGGAATAGATCGGGTTTATCGGGTGGATGGCCGGCCAAGGT 1320
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 AGTGCCCACTGGAATAGATCGGGTTTATCGGGTGGATGGCCGGCCAAGGT 1320

Qy 1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT 1380
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT 1380
```

```
Qy 1381 GTTTAGCAGTGTACACTACCTTCTGATCTCTGATCTACTACTGCCTCTCGGGGCAA 1440
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 GTTTAGCAGTGTACACTACCTTCTGATCTCTGATCTACTACTGCCTCTCGGGGCAA 1440

Qy 1441 GTGGTATCAGTTGCTCCTAGCGGTTATGTCTTTACATAGAAGGGACCTCTGGAAACCCCAT 1500
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 GTGGTATCAGTTGCTCCTAGCGGTTATGTCTTTACATAGAAGGGACCTCTGGAAACCCTAT 1500

Qy 1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCTTTGATGATACCATGTCC 1560
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCTTTGATGATACCATGTCC 1560

Qy 1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC 1620
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC 1620

Qy 1621 CAGGCCTATCACTCTAGAGTATAAACAACTCCATATCTTGGTACCCCTATACAATCCTGG 1680
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 CAGGCCTGTCACTCGAGTATAAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG 1680

Qy 1681 TGCGAGGGGATGTATGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA 1740
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 TGCGAGGGGATGTATGTTAAATTCAAAAATAACACATGGGGTTGCCGTATTCGCAA 1740

Qy 1741 TGTGCCATCGTACTGCACTATGGGCACTGATGCCAGTGTGGAACGACACTCGCAACACTTA 1800
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 TGTGCCATCGTACTGCACTATGGGCACTGATGCCAGTGTGGAACGACACTCGCAACACTTA 1800

Qy 1801 CGAAGCATGCGGTGTAACACCATGGCTAACAACCGCACAACGGCTCAGCCCTGAA 1860
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 CGAAGCATGCGGTGTAACACCATGGCTAACAACCGCACAACGGCTCAGCCCTGAA 1860

Qy 1861 ATTGGCTATATACAATACCCTGGGTCTAAAGAAAATGTTTAAACCTCATAATTGGATGTC 1920
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 ATTGGCTATATACAATACCCTGGGTCTAAAGAAAATGTTTAAACCTCATAATTGGATGTC 1920

Qy 1921 AGGCCATTTGTATTTGAGGGATCAGATCAGATACCCTATAGTTTACTTTATGACCCTGTGAA 1980
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921 AGGCCATTTGTATTTGAGGGATCAGATCAGATACCCTATAGTTTACTTTATGACCCTGTGAA 1980

Qy 1981 TTCCACTCTCCTACCACCGGAGAGGTGGCTAGGTTGCCCGGTACCCACCTGTGGTACG 2040
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981 TTCCACTCTCCTACCACCGGAGAGGTGGCTAGGTTGCCCGGTACCCACCTGTGGTACG 2040

Qy 2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAAGACCTAGCCACAGG 2100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAAGACCTAGCCACAGG 2100

Qy 2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGTCTATATATTCCGCACGCGGTGC 2160
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGTCTATATATTCCGCACGCGGTGC 2160

Qy 2161 TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGTCTAATTCTGTCCCCTTTGTTGGGGTGTGG 2220
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161 TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGTCTAATTCTGTCCCCTTTGTTGGGGTGTGG 2220

Qy 2221 CAGCAAGTATCTTATTTTAGCCTGTCGTCCATCCGTCGTATCGTGTACCCCAGTCTGGATGT 2280
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221 CAGCAAGTATCTTATTTTAGCCTGTCGTCCATCCGTCGTATCGTGTACCCCAGTCTGGATGT 2280

Qy 2281 TGGTTACCCTTTGCGTCCATGTCGTGTGCTCCATCCGTGCTTCTTCATCTCCAAGCTGGGGATCTTC 2340
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2281 TGGTTACCCTTTGCGTCCATGTCGTGTGCTCCATCCGTGCTTCTTCATCTCCAAGCTGGGGATCTTC 2340

Qy 2341 TTTGTCTAAAGCTCAAGTAGCTCCTTTTGATTTTCTTCATCTGTGTCATCTCCG 2400
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2341 TTTGTCTAAAGCTCAAGTAGCTCCTTTTGATTTTCTTCATCTGTGTCATCTCCG 2400

Qy 2401 CTGCAGGCTACGTTATGCTGCCCATGGCTTGTGTGCCCATGGCTTGCCCCT 2460
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401 CTGCAGGCTACGTTATGCTGCCCATGGCTTGTGTGCCCATGGCTTGCCCCT 2460
```

```
Qy  2461  AACTTTCTTGTTGCAGCAGCTGCTGCCCAACCAGATTATGATGACTGGTGCGACTGCT  2520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  AACTTTCTTGTTGCAGCAGCTGCTGCCCAACCAGATTATGATGACTGGTGCGACTGCT  2520

Qy  2521  AGTGGCAGGGTAGTTGTTGTGGGCCGGCCGTAACCGTGGTCACCGTAGCTTCGCTTGT  2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  AGTGGCAGGGTAGTTGTTGTGGGCCGGCCGTAACCGTGGTCACCGTAGCTTCTGCTTGT  2580

Qy  2581  AGGTCCTTCGGCCTCGGTAGCCTTCGGTCTTTAACCCTTGGTTTACGCCTTGGTTTACGCCTGCTTCAGC  2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  AGGTCCTTCGGCCTCGGTAGCCTTCGGTCTTTAACCCTTGGCTCCACGCCTGCTTCAGC  2640

Qy  2641  TTTTGATACCGGAGATAATTGGAGGGCTGACATATACCACCTGTAGTAGCATTAGTTGTCAT  2700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  TTTTGACACCGGAGATAATTGGAGGGCTGACATATACCACCTGTAGTAGCATTAGTTGTCAT  2700

Qy  2701  GTCTCGTTTGGCTTCTTGCTCACTGTGCTGTGTGCTTTAGTTAACTCCTATCT  2760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  GTCTCGTTTGGCTTCTTGCTCACTGTGCTGTGTGCTTTAGTTAACTCCTATCT  2760

Qy  2761  TTGGCAACGTTGGGAGAATTGGTTTGGAACGTTACACTAAGACCGGAGAGGTTTCCT  2820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  TTGGCAACGTTGGGAGAATTGGTTTGGAACGTTACACTAAGACCGGAGAGGTTTCTCCT  2820

Qy  2821  TGTGCTGGTTGTTTCCCCGGTGCGACATATGACGCGCTGGTGACTTTCTGTGTGTCA  2880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  TGTGCTGGTTGTTTCCCCGGTGCGACATATGACGCGCTGGTGACTTTCTGTGTGTCA  2880

Qy  2881  CGTAGCTCCTTCTATGTTTAACATCCAGTGCAGACATCGTTCTTTGGGACTGACTCTAGGGT  2940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  CGTAGCTCCTTCTATGTTTAACATCCAGTGCAGACATCGTTCTTTGGGACTGACTCTAGGGT  2940

Qy  2941  TAGGGCCCATAGAATGTTGGGTCCTCGGAAAGTGTCATGCTTGGTATTCTCATTATGT  3000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941  TAGGGCCCATAGAATGTTGGGTCGTCTCCGGAAAGTGTCATGCTTGGTATTCTCATTATGT  3000

Qy  3001  TCTTAAGTTTTTCCTCTAGTGTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA  3060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  TCTTAAGTTTTTCCTCTAGTGTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA  3060

Qy  3061  TGGTGATGTCTTGCCTAATGATTTTGCCTGAAAACTACCGACCTTTTACGCTGAACGTGG  3120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061  TGGTGATGTCTTGCCTAATGATTTTGCCTGAAAACTACCGACCTTTTACGCTGAACGTGG  3120

Qy  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAGACGCTCGGCGTGTGGGGACAC  3180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAGACGCTCGGCGTGTGGGGACAC  3180

Qy  3181  GGTTGATGGTTTGCCCGTTGTTGCCGGACTTGTTTCGCAGGGTTGGCTAT  3240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181  GGTTGATGGTTTGCCCGTTGTTGCCGGACTTGTTTCGCAGGGGTTAGCTAT  3240

Qy  3241  GCCGCCAGATGGTGGGCCATTACCGCACCTTTTACGGTGTCTCTGACGTGG  3300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241  GCCGCCAGATGGTGGGCCATTACCGCACCTTTTACGGTGTCTCTGACGTGG  3300

Qy  3301  CACGCGTGTCAGCGATGGCAGCGCTAATGCACATCTCATCACCCCCGACTTGGACTGGAAC  3360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301  CACGCGCTGTCAGCGATGGCAGCGCTAATGCACGCGGACATCTATCAACCACCCGACTTGGACTGGAAC  3360

Qy  3361  TATCTTCAGATTAGGATCTCTGGCCACTAGCTACATGGGATTTGTTTGTGACAACGTGTT  3420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361  TATCTTCAGATTAGGATCTCTGGCCACTAGCTACATGGGATTTGTTTGTGACAACGTGTT  3420

Qy  3421  GTATACTGCTCACATGGCAGCAAGGGGCGCCGGTTGGCTCATCCCACAGGCTCTATACA  3480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421  GTATACTGCTCACATGGCAGCAAGGGGCGCCGGTTGGCTCATCCCACAGGCTCCATACA  3480

Qy  3481  CCCAATAACCGTTGACGCGCTAATGACGCGGCTAATCTTATCAACCACCATGTGGAGCTGG  3540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481  CCCAATAACCGTTGACGCGCTAATGACGACATCTATCAACCACCATGTGGAGCTGG  3540

Qy  3541  GTCCCTTACTCGGTGCTCTTGCGGGGGAGACCAAGGGGTATCTCGGTAACGACGACTGGGGTC  3600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy  3541  GTCCCTTACTCGGTGCTCTTGCGGGGGAGACCAAGGGGTATCTCGGTAACGACGACTGGGGTC  3600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541  GTCCCTTACTCGGTGCTCTTGCGGGGGAGACCAAGGGGTATCTCGGTAACGACGACCGCCCCAT  3600

Qy  3601  ATTGGTTGAGGTCAACAAATCCGATGGTGTCAATGTGTGTGCGGGGCCCTTCCCAT  3660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601  ATTGGTTGAGGTCAACAAATCCGATGGTGTCAATGTGTGTGCGGGGCCCTTCCCAT  3660

Qy  3661  GGCTGTTGCCAAGGGTTCTTCAGTGCCGATTCTGTGCTCCTCCCGGGCATGTTATTGG  3720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661  GGCTGTTGCCAAGGGTTCTTCAGTGCCGATTCTGTGCTCCTCCCGGGCATGTTATTGG  3720

Qy  3721  GATGTTCACCGCTGTAGAAATTCTGGCGGTTCAGTGTTCAGTAGAGGGTTAGGCCGTT  3780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721  GATGTTCACCGCTGTAGAAATTCTGGCGGTTCAGTGTTCAGTCAGCCAGATTAGGGTTAGGCCGTT  3780

Qy  3781  GGTGTGTGCTGGATACCATCCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC  3840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781  GGTGTGTGCTGGATACCATCCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC  3840

Qy  3841  TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC  3900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841  TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC  3900

Qy  3901  CAAATTACCACTTTCTTACGATGCAGGAGAGTATGAGGGTTCTTGTCCTAAATCCCAGTGT  3960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901  CAAATTACCACTTTCTTACGATGCAGGAGAGTATGAGGGTTCTTGGTCCTAAATCCCAGTGT  3960

Qy  3961  GGCTACAACAGCATCAATGCCAAAGTACATGCCACGCGACGTACGGCGGTGAATCCAAATTG  4020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961  GGCTACAACAGCATCAATGCCAAAGTACATGCCACGCGACGTACGGCGGTGAATCCAAATTG  4020

Qy  4021  CTATTTTAATGGCAAATGTACCAACACAACAGGGGCTTCACTTACGTACACGACATATGGCAT  4080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021  CTATTTTAATGGCAAATGTACCAACACAACAGGGGCTTCACTTACGTACACGACATATGGCAT  4080

Qy  4081  GTACCTGACCGGAGCATGTTCCCGGACTGGAACTAATCATTGTGACGAATGCCATGC  4140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081  GTACCGACCGGAGCATGTTCCCGGACTGGAACTAATCATCATTGTGACGAATGCCATGC  4140

Qy  4141  TACCGATGCAACCACCGTGTTGGGCATTGGAAAGGTCCATCCGAAGCTCCATCCAAAAA  4200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141  TACCGATGCAACCACCGTGTTGGGCATTGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA  4200

Qy  4201  TGTTAGGCTAGTGTCTTCTTGCCACACGGCTACCCCCCCTGAGTAATCCTACCACCACATGC  4260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201  TGTTAGGCTAGTGTCTTCTTGCCACACGGCTACCCCCCCTGAGTAATCCTACCACACATGC  4260

Qy  4261  CAACATAACTGAGATTCAATTAACCGATGAAGGCACTGATCCCCTTCATGGAAAAAGAT  4320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261  CAACATAACTGAGATTCAATTAACCGATGAAGGCACTGATCCCCTTTCATGGAAAAAAGAT  4320

Qy  4321  TAAGGAGGAAAATCTGAACGAAAAGGGAGACACCTTATCTTGAGGCTACCAAAACACTG  4380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321  TAAGGAGGAAAATCTGAACGAAAAGGGAGACACCTTATCTTGAGGCTACCAAAACACTG  4380

Qy  4381  TGATGAGCTTGCTAACGAGTTAGCTCGAACCCCTGAGGGGCACTGTGTAGTTGCCACTGG  4440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381  TGATGAGCTTGCTAACGAGTTAGCTCGAACCCCTGAGGGGCACTGTGTAGTTGCCACTGG  4440

Qy  4441  ATGTGACATCTCAAAAATGTTGTTTGTTGTGACAACGTGTT  4500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441  ATGTGACATCTCAAAAATGTTGTTTGTTGTGACAACGTGTT  4500

Qy  4501  TACAGGGTACACTGGTGTATGACTGCAGCCTCATCATGGTAGAAGGCAC  4560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501  TACAGGGTACACTGGTGTATGACTGCAGCCTCATCATGGTAGAAGGCAC  4560

Qy  4561  ATGCCATGTTGACCTTGACCCTGTTCGTTGTGTCGGGGTTTCAGC  4620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561  ATGCCATGTTGACCTTGACCCTGTTCGTTGTGTCGGGGTCTCAGC  4620

Qy  4621  AATAGTTAAAGGCCAGCCGCCGCCACAGGCCGTGGGAGAGCCTGGGGGTC  4680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Qy   5761 TGCTGGTTAGTCTACTTCCGGCGTTCAATCTCAATGAGAGTTGTGGGCGTCTTGTCAGC 5820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5761 TGCTGGTTAGTCTACTTCCGGCGTTCAATCTCAATGAGAGTTGTGGGCGTCTTGTCAGC 5820

Qy   5821 TTGTGCAATGTTGCTTGACAACAGCAGGCCAGATCACTGCCCAACAGACTTCTTAC   5880
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5821 TTGTGCAATGTTGCTTGACAACAGCAGGCCAGATCACTGCCCAACAGACTTCTTAC   5880

Qy   5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940

Qy   5941 CAGGAAGATACTGGGCATTCTGGAGCATCCACCCCTGGAGTGTCATATCAGCTTGCAT   6000
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5941 CAGGAAGATACTGGGCATTCTGGAGCATCCACCCCTGGAGTGTCATATCAGCTTGCAT   6000

Qy   6001 CCGTTGGCTCCACACCCCGACGGAGATGATTGCGGCCTCATTGGGGTCTAGAGAT   6060
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6001 CCGTTGGCTCCACACCCCGACGGAGATGATTGCGGCCTCATTGGGGTCTAGAGAT   6060

Qy   6061 TTTGGCAGTATGTGTGCAATTTCTTTGTGATTTGTGTATTTGTCTTTAAAGCTGGAGTTCA 6120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6061 TTTGGCAGTATGTGTGCAATTTCTTTGTGATTTGTGTATTTGTCTTTAAAGCTGGAGTTCA 6120

Qy   6121 GAGCATGTTAACATTCCTGGTTGTCCTTCTTCACAGCTGCCAGAAGGGGTACAAGGGCCC 6180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6121 GAGCATGTTAACATTCCTGGTTGTCCTTCTTCACAGCTGCCAGAAGGGGTACAAGGGCCC 6180

Qy   6181 CTGGATTGGATCAGGTATGCTCCAAGACGCGTGTCCATGCGGTGCTGAACTCATCTTTTC 6240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6181 CTGGATTGGATCAGGTATGCTCCAAGACGCGTGTCCATGCGGTGCTGAACTCATCTTTTC 6240

Qy   6241 TGTTGAGAATGGTTTTGCAAAACTTTACAAGGACCCAGAACTTGTTCAAATTACTGGAG   6300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6241 TGTTGAGAATGGTTTTGCAAAACTTTACAAGGACCCAGAACTTGTTCAAATTACTGGAG   6300

Qy   6301 AGGGGCTGTTCCAGTCAACGCTAGGCTGTGTGGGTCGGCTAGAACCGACCAACTGATTG   6360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6301 AGGGGCTGTTCCAGTCAACGCTAGGCTGTGTGGGTCGGCTAGAACCGACCAACTGATTG   6360

Qy   6361 GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA 6420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6361 GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA 6420

Qy   6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCAAAC   6480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCAAAC   6480

Qy   6481 CTTGAGAGCTGCAGTGCCGTGGCCGTGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6481 CTTGAGAGCTGCAGTGCCGTGGCCGTGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540

Qy   6541 AACTCCCTGGACGACATCGCTTGCTGCTTGACAAGCTGAAAAACTGTTAA   6600
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6541 AACTCCCTGGACGACATCGCTTGCTGCTTGACAAGCTGAAAAACTGTTAA   6600

Qy   6601 GCTTCCCTTCCGCGTTGACGGTCACACCTGGTGTGCGCATGCAACTTAATTGCGTGA   6660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6601 GCTTCCCTTCCGCGTTGACGGTCACACCTGGTGTGCGCATGCAACTTAATTGCGTGA   6660

Qy   6661 TGCACTTGAGACAAATGACTGTAATTCCACAAACACTCCTAGTGTGATGAGAGCCGCAGT   6720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6661 TGCACTTGAGACAAATGACTGTAATTCCATAAACACTCCTAGTGTGATGAGAGCCGCAGT   6720

Qy   6721 GTCCGCTCTTGTTTCAAACAGAGTTGCGGCTACAAACCAATTGCTTGAGGCAATTTC   6780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6721 GTCCGCTCTTGTTTCAAACAGAGTTGCGGCTACAAACCAATTGCTTGAGGCAATTTC   6780

Qy   6781 AGCTGGCGTTGACACCACCAAACTGCCACCCCTCCATCGAAGAGGTAGTGTTAAGAAA   6840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6781 AGCTGGCGTTGACACCACCAAACTGCCACCCCTCCATCGAAGAGGTAGTGTTAAGAAA   6840

Db   4621 AATAGTTAAAGGCCAGCTGGAGGGGCCGCACAGGCCGTGGGAGAGCTGGCATATACTACTA 4680

Qy   4681 TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTGTTCCTGAATGCAACATTGTTGAAGCCTT 4740
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4681 TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTGTTCCTGAATGCAACATTGTTGAAGCCTT 4740

Qy   4741 CGACCAGCCAAGGCATGGTATGGTTGTCATCAACAGAGCTCAAACTATTCTGACACAC   4800
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 CGACCAGCCAAGGCATGGTATGGTTGTCATCAACAGAGCTCAAACTATTCTGGACACAC   4800

Qy   4801 CTATCGCACCCAACCTGGGTTACCTGCAGATAGGAGCAAATTGGACGAGTGGGCTGATCT 4860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 CTATCGCACCCAACCTGGGTTACCTGCAGATAGGAGCAAATTGGACGAGTGGGCTGATCT 4860

Qy   4861 CTTTTCTATGGTCAACCCGAACCTTCATTTGTCAATACTGCAAAAAGAACTGCTGACAA   4920
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861 CTTTTCTATGGTCAACCCGAACCTTCATTTGTCAATACTGCAAAAAGAACTGCTGACAA   4920

Qy   4921 TTATGTTTTGTGACTGCAGTGCGCCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCC 4980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 TTATGTTTTGTGACTGCAGTGCGCCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCC 4980

Qy   4981 CAATGACGACCACGGTGGCAGGAGCCCGGCTGGCAGGGAGCCCGGCTTGGGGAAAAAAACCTTGTGGGGGTTCTGTG 5040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 CAATGACGACCACGGTGGCAGGAGCCCGGCTGGCAGGGAGCCCGGCTTGGGGAAAAAAACCTTGTGGGGGTTCTGTG 5040

Qy   5041 GCGCTTGGACGGCGCGTCACGGCCTGACGGCGCTGACCGCGGGAGCGCTGACGGCTGAGTGCGCTGACCAGATACCA 5100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041 GCGCTTGGACGGCGCGTCACGGCCTGACGGCGCTGACCGCGGGAGCGCTGACGGCTGAGTGCGCTGACCAGATACCA 5100

Qy   5101 AATGTGCTTCACTGAGTCAATATCATCATTCATTCATTCTCGGGACAGCGCCGCACTCGGCTGTCGTGGCGTTGGAGT 5160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101 AATGTGCTTCACTGAGTCAATATCATCATTCATTCATTCTCGGGACAGCGCCGCACTCGGCTGTCGTGGCGTTGGAGT 5160

Qy   5161 GGCTATCTTCAGCTTTATCTAGCCATTGACACTTTTGGGCCACTTTTGGCGCCACTTGTGTGCGCCGGTTGCGCGGTC 5220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161 GGCTATCTTCAGCTTTATCTAGCCATTGACACTTTTGGGCCACTTTTGGCGCCACTTGTGTGCGCCGGTTGCGCGGTC 5220

Qy   5221 TATTACATCAGTCCCTACCGGTCTACTGTCCCCCCAGTGGTTGACGAAGAGAAATCGT   5280
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221 TATTACATCAGTCCCTACCGGTCTACTGTCCCCCCAGTGGTTGACGAAGAGAAATCGT   5280

Qy   5281 GGAGGAGTGTGCATCAATCATTCATTCCCTTGGAGGAGCCGTTGCTGCGATTGACAAGCTGAA 5340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5281 GGAGGAGTGTGCATCAATCATTCATTCCCTTGGAGGAGCCGTTGCTGCGATTGACAAGCTGAA 5340

Qy   5341 CAGTACAATCACCACACAGTCCTTTGCATCATGCGTGTTTGAAAACCCCCTTGAAAACTTAACTAC 5400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5341 GAGTACAATAACCACAACAGTCCTTTGCATCATGCGTGTTTGAAAACCGCCCTTGAAAACTTAACTAC 5400

Qy   5401 CTTTCTTGGGGCTCATGCAGGCTACATCCTTGTCATCGCTACTAGAGTATTGCTGTGGTTTAGT 5460
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   5401 CTTTCTTGGGGCTCATGCAGGCTACATCCTTGTCATCGCTACTAGAGTATTGCTGTGGTTTAGT 5460

Qy   5461 CACTTTACCTACCTGACAATCCCCTTTTGCATCATGCGTGTTTGCTTTTCATTGCGGGTATTACTAC 5520
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5461 CACTTTACCTACCTGACAATCCCCTTTTGCATCATGCGTGTTTGCTTTTCATTGCGGGTATTACTAC 5520

Qy   5521 CCCACTACCTTCACAGAGATCAAAAGTGTTCCTGTTATTTGGAGGCGCAATTGCGTCCAA 5580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5521 CCCACTACCTTCACAGAGATCAAAAGTGTTCCTGTTATTTGGAGGCGCAATTGCGTCCAA 5580

Qy   5581 GCTTACAGAACGCGTAGAGCGCCACTCGGCGGTTCTTCATGATGGCCGGGAACAGCTCT   5640
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5581 GCTTACAGAACGCGTAGAGCGCCACTCGGCGGTTCTTCATGATGGCCGGGAACAGCTCT   5640

Qy   5641 TGGTACATGGACATCCGTGGGTTTTGTTCTTTGACATGCTTAGGCGGCTCATGCTGCCGCCCTC 5700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5641 TGGTACATGGACATCCGTGGGTTTTGTTCTTTGACATGCTTAGGCGGCTCATGCTGCCCCTC 5700

Qy   5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGAGGTGAGTGGAGTGGCCACTATGGGATCAGCT 5760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGGCCACTATGGGATCAGCT 5760

Qy 6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCTCCGAGATCCGTCCC 6900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCTCCGAGATCCGTCCC 6900

Qy 6901 AGGAGTGTCATGTCCTGAAAGCTGCAACGAAGTGACCCGTTAGAAGGTCCTTCAAACCT 6960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6901 AGGAGTGTCATGTCCTGAAAGCTGCAACGAAGTGACCCGTTAGAAGGTCCTTCAAACCT 6960

Qy 6961 CCCTTCCTTCACCACTGTTCTACACAGTTGGCCATGCGGATCGCCGATGCGGGTGA 7020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6961 CCCTTCCTTCACCACTGTTCTACACAGTTGGCCATGCGGATCGCCGATGCGGGTGA 7020

Qy 7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCCGAAACAGGCGGAGGCCCTGATGA 7080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCCGAAACAGGCGGAGGCCCTGATGA 7080

Qy 7081 TTTACCCAGTTACCCTCCCAAAAATCAGCGCCCGCCCAAACGGCCTACAACCGCTTCAGCT 7140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7081 TTTACCCAGTTACCCTCCCAAAAATCAGCGCCCGCCCAAACGGCCTACAACCGCTTCAGCT 7140

Qy 7141 GGCTACACACCGCTTCCAGCTACTGCTTACTCGTTACTCGGCCTTAAGATACGGGGAAAGGA 7200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7141 GACTACAACCGCTTCCAGCTACTGCTTACTCGTTACTCGGCCTTAAGATACGGGGAAAGGA 7200

Qy 7201 TTCCACTTCAGTCAGCATGAGCTACACCTGGACCGACGGTGATTAGCTTCAAACTGCTTCTAAAGT 7260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7201 TTCCACTTCAGTCAGCATGAGCTACACCTGGACCGACGGTGATTAGCTTCAAACTGCTTCTAAAGT 7260

Qy 7261 TTCGTGAGCATGAGCTACACCTGGACCGACGGTGATTAGCTTCAAACTGCTTCAAAAGT 7320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7261 TTCGTGAGCATGAGCTACACCTGGACCGACGGTGATTAGCTTCAAACTGCTTCAAAAGT 7320

Qy 7321 TCTGTCTGCAACTCGGGCCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGT 7380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7321 TCTGTCTGCAACTCGGGCCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGT 7380

Qy 7381 GACTGAGCCGCGGGATGCGGAGCTTAGAAACAAAAGTCACTATTAATAGACAACCTCT 7440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7381 GACTGAGCCGCGGGATGCGGAGCTTAGAAACAAAAGTCACTATTAATAGACAACCTCT 7440

Qy 7441 GTTCCCCCATCATACCACCAACAAGCAAGTGAGATTGGCTAAGGCAAAAGCTTCAAAGTTGT 7500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7441 GTTCCCCCATCATACCACCAACAAGCAAGTGAGATTGGCTAAGGCAAAAGCTTCAAAGTTGT 7500

Qy 7501 CGGTGTCATGTGGGACTATGATGAAGTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTC 7560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7501 CGGTGTCATGTGGGACTATGATGAAGTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTC 7560

Qy 7561 CCACATCACTGGCCTTCGGGCACTGATGTCGTTCTCGGAGCAGCCCAAGGCTGTTCT 7620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7561 CCACATCACTGGCCTTCGGGCACTGATGTCGTTCTCGGAGCAGCCCAAGGCTGTTCT 7620

Qy 7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680

Qy 7681 AGTTCCAAAGGAGGAGGTCTTCGTGAGCAGACCCCCAGAGAAACCCCCAAG 7740
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7681 AGTTCCAAAGGAGGAGGTCTTCGTGAGCAGACCCCCAGAGAAACCCCCAAG 7740

Qy 7741 GCTTATCTCGTACCCCCACCTTGCATGCGTACGGTACGGTACGGTCAGGT 7800
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7741 GCTTATCTCGTACCCCCACCTTGCATGCGTACGGTACGGTACGGTCAGGT 7800

Qy 7801 TGCTCCTGACGTAGTTAAAGCTGTCTGAGGCAGATGCGTTTGTAGATCCACGTAC 7860
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7801 TGCTCCTGACGTAGTTAAAGCTGTCTGAGGCAGATGCGTTTGTAGATCCACGTAC 7860

Qy 7861 CCGTGTCAAGCGCTCGTTGTCGATGTGTCGATGGTGTCACCGAGCCACATGCGGATAC 7920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7861 CCGTGTCAAGCGCTCGTTGTCGATGTGTCGATGGTGTCACCGAGCCACATGCGGATAC 7920

Qy 7921 AGTGTGTTTGACAGTGACTACCATCACACCCGAGGATATCATGGTCGGAGACAGATCTACTC 7980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7921 AGTGTGTTTGACAGTGACTACCATCACACCCGAGGATATCATGGTCGGAGACAGATCTACTC 7980

Qy 7981 AGCAGCTAAACTCAGTGACCAACACCGCATTCACGAGCTGGCATTGCGGAGGCAGTTATA 8040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7981 AGCAGCTAAACTCAGTGACCAACACCGCATTCACGAGCTGGCATTGCGGAGGCAGTTATA 8040

Qy 8041 CGCTGGAGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGGAGTGTAGGTC 8100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8041 CGCTGGAGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGGAGTGTAGGTC 8100

Qy 8101 TTCCGGCGTCTATATACTACCTCAAGTTGACCTGCTGGCTGAAGGTAAATGC 8160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8101 TTCCGGCGTCTATATACTACCTCAAGTTGACCTGCTGGCTGAAGGTAAATGC 8160

Qy 8161 TGCAGCCGAACAGGCTGGCATGGCGGCGCCGAGCGATGAGAACCCCGCTTCCTTATTGCGGCGATGATTGCAC 8220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8161 TGCAGCCGAACAGGCTGGCATGGCGGCGCCGAGCGATGAGAACCCCGCTTCCTTATTGCGGCGATGATTGCAC 8220

Qy 8221 CGTAATTTGGAAGAGCGCCGAGCAGCAGATGCAGACAAGCAATGCGTGTCTTTGCTAG 8280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8221 CGTAATTTGGAAGAGCGCCGAGCAGCAGATGCAGACAAGCAATGCGTGTCTTTGCTAG 8280

Qy 8281 CTGGATGAAGGTGATGGGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8281 CTGGATGAAGGTGATGGGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340

Qy 8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400

Qy 8401 CTACTTCTTACAAGAGATCCTGGTATCCCCCTTGGCAGTGTCTCGCCGAGGTCTCGGG 8460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8401 CTACTTCTTACAAGAGATCCTGGTATCCCCCTTGGCAGTGTCTCGCCGAGGTCTCGGG 8460

Qy 8461 ATACAACCCAGTGCTGCGGTGGATTGGGTATCTAATACTCAGTTGTTGTGGGGT 8520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8461 ATACAACCCAGTGCTGCGGTGGATTGGGTATCTAATACTCAGTTGTTGTGGGGT 8520

Qy 8521 TAGCCGTGTGTGGCTGTCCATTTCATGGGCAGCAGATGTCTCTTTGGAGGACAAACTTCCCGA 8580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8521 TAGCCGTGTGTGGCTGTCCATTTCATGGGCAGCAGATGTCTCTTTGGAGGACAAACTTCCCGA 8580

Qy 8581 GACTGCAGCCTTTGACTGTGATGGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8581 GACTGCAGCCTTTGACTGTGATGGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640

Qy 8641 CATCATTGCTGGTGTGTGCACGGTATTGGGCTTTCTCGGGGTGCGGTGGCTACCAACGCTGA 8700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8641 CATCATTGCTGGTGTGTGCACGGTATTGGGCTTTCTCGGGGTGCGGTGGCTACCAACGCTGA 8700

Qy 8701 GATCCTCCAGAGTTTCCCAATCACTCAACAGACATGCCGAGCCCCCTGCGAGCCTGGCG 8760
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8701 GATCCTCCAGAGTTTCCCAATCACTCAACAGACATGCCGAGCCCCCTGCGAGCCTGGGCG 8760

Qy 8761 AAAGAAAGCCAGGCGCCAAGAGCGCCAGCGCCGGAGCACACGCAAATT 8820
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 8761 AAAGAAAGCCAGGCGCCAAGAGCGCCAGCGCCGGAGCACACGCAAATT 8820

Qy 8821 GGCTCGCTTCCTTCTGCATGCTACATCTAGACCTCTACACTTTGATATGAGACGAG 8880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8821 GGCTCGCTTCCTTCTGCATGCTACATCTAGACCTCTACACTTTGATAAGACGAG 8880

Qy 8881 CGTTGGCTCCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGGAGGGGATGTGTT 8940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8881 CGTTGGCTCCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGGAGGGGATGTGTT 8940

Qy 8941 TATTACACCACAGAAGAATTGCAGAAGTTCCTTGTGAAGTATTGGCTGTCATTGTTTT 9000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8941 TGTTACACCACAGAAGAATTGCAGAAGTTCCTTGTGAAGTATTGGCTGTCATTGTTTT 9000

Qy 9001 TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db   9001 TGCCCTAGGGCTCATTGCTGTTGGACTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA 9060
          ||||||
Qy   9061 CTAACAG---TTTTTTTTTTTTTTTTTTTTTTAGGCAGCGCAACAGGGGAGACCCC 9116
                       ||||||||||||||||||||||||||||||
Db   9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTAGGCAGCGCAACAGGGGAGACCCC 9120

Qy   9117 GGGCTTAACGACCCCGC 9133
          |||||||||||||||||
Db   9121 GGGCTTAACGACCCCGC 9137
```

RESULT 10
AR310126
LOCUS   AR310126   9143 bp   DNA   linear   PAT 12-JUN-2003
DEFINITION Sequence 390 from patent US 6558898.
ACCESSION   AR310126
VERSION   AR310126.1 GI:31702404
KEYWORDS   .
SOURCE   Unknown.
ORGANISM   Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 9143)
AUTHORS   Simons,J.N., Piloc-Matias,T.J., Dawson,G.J., Schlauder,G.G.,
            Desai,S.M., Leary,T.P., Muerhoff,A.S., Erker,J.C., Buijk,S.L. and
            Mushahwar,I.K.
TITLE   Non-A, non-B, non-C, non-D, non-E hepatitis reagents and methods
            for their use
JOURNAL   Patent: US 6558898-A 390 06-MAY-2003;
FEATURES   Location/Qualifiers
     source   1..9143
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match   96.4%; Score 9059.8; DB 6; Length 9143;
Best Local Similarity   99.6%;
Matches 9096; Conservative   0; Mismatches   0; Indels 37; Gaps 1;

```
Qy      1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGAGCACCCCCCTAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGAGCACCCCCCTAG 60

Qy     61 CAGGGCGTGGGGGATTTCCCCGCCCTCTGCAGAAGGGTGGAGCGAGGTCGGATGGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAGGGCGTGGGGGATTTCCCCGCCCTCTGCAGAAGGGTGGAGCGAGGTCGGATGGC 120

Qy    121 GTAGGCGCGGGGGACTCATGACGGCCGCGCGCGACTGATGAGCAAGCGCCAAGTTGACTTGGATGGC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTAGGCGCGGGGGACTCATGACGGCCGCGCGCGACTGATGAGCAAGCGCCAAGTTGACTTGGATGGC 180

Qy    181 CCTGATGGGCGTTCATGGGTTCGGTGGTGGTGGGTGTGGTGGGCGCGTTAGGCAGCGCCAAGTTGACTTGGATGGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCTGATGGGCGTTCATGGGTTCGGTGGTGGTGGGTGTGGTGGGCGCGTTAGGCAGCGCCAAGTTGACTTGGATGGC 240

Qy    241 CCTCCCAGATAGAGCGGCGGCGGCACTGTAGGGAAGAACCGGGGACCGGTCACTACCAAGGACG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTCCCAGATAGAGCGGCGGCGGCACTGTAGGGAAGAACCGGGGACCGGTCACTACCAAGGACG 300

Qy    301 CAGACCTCTTTTGAGTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACTATATGTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGACCTCTTTTGAGTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACTATATGTGT 360

Qy    361 TGGGATGGTTGGGGTTAGCCATCCATCCATCCGTACCGTAGGGTCCTTGCGAGGGGAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TGGGATGGTTGGGGTTAGCCATCCATCCATCCGTACCGTAGGGTCCTTGCGAGGGGAT 420

Qy    421 CTGGGAGTCTCGTAGACCGTAGCCTTGTATTTCTACTCAAACAAGTCCTGTACC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGGGAGTCTCGTAGACCGTAGCCTTGTATTTCTACTCAAACAAGTCCTGTACC 480

Qy    481 TGCGCCCAGAACGCGCAAGAACAAGCAGACAGACGCAGGCTTCATATCCTGTGTCATTAAAAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGCGCCCAGAACGCGCAAGAACAAGCAGACAGACGCAGGCTTCATATCCTGTGTCATTAAAAC 540
```

```
Qy    541 ATCTGTTGAAAGGGACAACGAGCAAAGCGCAAAGTCCAGCGCGATGCTCGGCTCGTGTAA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ATCTGTTGAAAGGGACAACGAGCAAAGCGCAAAGTCCAGCGCGATGCTCGGCTCGTGTAA 600

Qy    601 TTACAAAATTGCTGGTATCCATGATGGCTTCGAGACATTGGCTCAGCGTCTTGCCAGC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTACAAAATTGCTGGTATCCATGATGGCTTCGAGACATTGGCTCAGCGTCTTGCCAGC 660

Qy    661 TCATGGTTGGGGACGCCAAGACCCTCGCATAAGTCTCGCAATCTTGGAATCTTCTGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TCATGGTTGGGGACGCCAAGACCCTCGCATAAGTCTCGCAATCTTGGAATCCTTCTGGA 720

Qy    721 TTACCCTTGGGGTGGATTGGTGATGTTACAACTCACAACTCACACACCTCTAGTAGGCCCGCTGT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TTACCCTTGGGGTGGATTGGTGATGTTACAACTCACAACTCACACACCTCTAGTAGGCCCGCTGT 780

Qy    781 GGCAGGAGCGGTCGTTCGACCAGTCGCTCGCAGATAGTACGCTTGCTGGAGGATGGAGTCAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GGCAGGAGCGGTCGTTCGACCAGTCGCTCGCAGATAGTACGCTTGCTGGAGGATGGAGTCAA 840

Qy    841 CTGGGCTACTGGTTGGTTCGGTGTCACCTTTTTGTGTGGTATGTCTCGTATCTTGGCCTG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTGGGCTACTGGTTGGTTCGGTGTCACCTTTTTGTGTGGTATGTCTCGTATCTTGGCCTG 900

Qy    901 TCCCTGAGTGGGGGCGGGTCACTGACCAGACACAAATACCACCAATCCTGACCAATTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCCCTGAGTGGGGGCGGGTCACTGACCAGACACAAATACCACCAATCCTGACCAATTG 960

Qy    961 CTGCCAGGTAATCAGGTTATCATTGTTCTCCTTCCACTTGCTACACGAGCCTGGTTG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CTGCCAGGTAATCAGGTTATCATTGTTCTCCTTCCACTTGCTACACGAGCCTGGTTG 1020

Qy   1021 TGTGATCTGTGCGGGACGAGTGCTGTGGGGTTCCCGCCAATCCGTACATCTCACACCCTTCAA 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TGTGATCTGTGCGGGACGAGTGCTGTGGGGTTCCCGCCAATCCGTACATCTCACACCCTTCAA 1080

Qy   1081 TTGGACTGGCACGGACTCCTTCTTGGCTGAGTGCTGACCACATTGTTATGGGCGCTCTTGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TTGGACTGGCACGGACTCCTTCTTGGCTGAGTGCTGACCACATTGTTATGGGCGCTCTTGT 1140

Qy   1141 GACCTGTGACGCCCTTGACATTGGTGAGTGTGTGCGTGTATTAGTCGGTGACTG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTGTGACGCCCTTGACATTGGTGAGTGTGTGCGTGTATTAGTCGGTGACTG 1200

Qy   1201 GCTTGTCAGGCACTGGTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GCTTGTCAGGCACTGGTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGA 1260

Qy   1261 AGTGCCACTGGAATAGATCCTGGGTTCCTAGGGTTCACAAGTTATCGGGTGAGTGGCCGCAAGGT 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGTGCCACTGGAATAGATCCTGGGTTCCTAGGGTTCACAAGTTATCGGGTGAGTGGCCGCAAGGT 1320

Qy   1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT 1380

Qy   1381 GTTTAGCAGTGTACACCGTACCTGGCGCGTTGGCGCTTCACTAGGAAGCCCAT 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GTTTAGCAGTGTACACCGTACCTGGCGCGTTGGCGCTTCACTAGAAGCCCAT 1440

Qy   1441 GTGGTATCAGTTGCTCCTTATGCTTTACATAGCTGAGTTTTGCTCGCCTTTGATGATACCATGTCC 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GTGGTATCAGTTGCTCCTTATGCTTTACATAGCTGAGTTTTGCTCGCCTTTGATGATACCATGTCC 1500

Qy   1501 CAGGGTGCCCACTGGATGCTCAATAGCTCAGTGAGTTTTGCTCGCCTTTGATGATACCAAGTGAC 1560
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CAGGGTGCCCACTGGATGCTCAATAGCTCAGTGAGTTTTGCTCGCCTTTGATGATACCAAGTGAC 1560

Qy   1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAATGTGTACAGTGTCCAAAGTGGAC 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAATGTGTACAGTGTCCAAAGTGGAC 1620
```

```
Qy  1621 CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTG 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 CAGGCCTGTCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTG 1680

Qy  1681 TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA 1740

Qy  1741 TGTGCCATCGTACTGCACTATGGGCACTGATGGAGTGTGGAACGACACTCGCAACACTTA 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 TGTGCCATCGTACTGCACTATGGGCACTGATGGAGTGTGGAACGACACTCGCAACACTTA 1800

Qy  1801 CGAAGCACATGCGGTGTAACACCATGGCTAACACCGATGCGCACAACGGCTCAGCCCTGAA 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 CGAAGCACGTCGGTGTAACACCATGGCTAACACCGATGCGCACAACGGCTCAGCCCTGAA 1860

Qy  1861 ATTGGCTATATTACAATACCCTGGTCTAAAGAAAATGTTAAACCTCATAATTGGATGTC 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 ATTGGCTATATTACAATACCCTGGTCTAAAGAAAATGTTAAACCTCATAATTGGATGTC 1920

Qy  1921 AGGCCATTGTATTTTGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAA 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 AGGCCATTGTATTTTGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAA 1980

Qy  1981 TTCCACTTCCTCCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCACTCTGGTACG 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 TTCCACTTCCTCCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCACTCTGGTACG 2040

Qy  2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100

Qy  2101 ATTGATCACCAAAGACAAAGACCTGGAAAAAATTATCAGGTCTTATATTCCGCACGGGTGC 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 ATTGATCACCAAAGACAAAGACCTGGAAAAAATTATCAGGTCTTATATTCCGCACGGGTGC 2160

Qy  2161 TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGTGTCTAATTCGTTGTGTGTGGG 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGTGTCTAATTCGTTGTGTGTGGG 2220

Qy  2221 CAGCAAGTATCTTATTTAGCTCCTGTGCTACTTGTACTTGTCCCTTTGTTTTGTTTGGGCGCTTC 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 CAGCAAGTATCTTATTTAGCTCCTGTGCTACTTGTACTTGTCCCTTTGTTTTGTTTGGGCGCTTC 2280

Qy  2281 TGGTTACCCTTTGCCTGTGCTCCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGT 2340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 TGGTTACCCTTTGCCTGTGCTCCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGT 2340

Qy  2341 TTTGTCTAAAGCTCAAGTAGCTCCTTTTTGATTTTCTTCATCTGTCTATCTCCG 2400
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 TTTGTCTAAAGCTCAAGTAGCTCCTTTTTGATTTTCTTCATCTGTCTATCTCCG 2400

Qy  2401 CTGCAGGCTACGTTATGCTTTGCCCTTTTAGGGTTGTGCCCATGGCTTACGGCCTTCAGC 2460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 CTGCAGGCTACGTTATGCTTTGCCCTTTTAGGGTTGTGCCCATGGCTTACGGCCTTCAGC 2460

Qy  2461 AACTTTCTTTGTTGTCAGCAGCTGCTGCGCCAACAGATATGACGTGGTGGGCGACTGCT 2520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 AACTTTCTTTGTTGTCAGCAGCTGCTGCGCCAACAGATATGACGTGGTGGGCGACTGCT 2520

Qy  2521 AGTGGGCAGGGTTAGTTTTGTGGGGCCGGCCCGTAACCGTGGTCACCGCCATAGCTTCTGT 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521 AGTGGGCAGGGTTAGTTTTGTGGGGCCGGCCCGTAACCGTGGTCACCGTAGCTTCTGT 2580

Qy  2581 AGTTCCTTGGCCTCTCGGGTCGTTTAACCCTCTTACGGCGCTTCAGC 2640
         |||||||||||||||||||||||||||||||||||||||||||
Db  2581 AGTTCCTTGGCCTCTCGGGTCGTTTAACCCTCTTACGGCGCTTCAGC 2640

Qy  2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTTATTAGTTGTCAT 2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641 TTTTGACACCGAGATAATTGGAGGGCTGACAATACCACCTGTTATTAGTTGTCAT 2700

Qy  2701 GTCTCGTTTTGGCTTCTTTTGCTCACTTGTTACCTCGTGCTGTTTAGTTAACTCCTATCT 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 GTCTCGTTTTGGCTTCTTTTGCTCACTTGTTACCTCGTGCTGTTTAGTTAACTCCTATCT 2760

Qy  2761 TTGGCAACGTTGGGAGAATTGGTTTGGAACGTTACACTGGTTACCTCCGGAGAGGTTTCCT 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 TTGGCAACGTTGGGAGAATTGGTTTGGAACGTTACACTGGTTACCTCCGGAGAGGTTTCCT 2820

Qy  2821 TGTGCTCGGTTGTTCCCCGTGCGACATATGACGCGTCGGTGTGACTTTCTGTGTCTGTGTCA 2880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821 TGTGCTCGGTTGTTCCCCGTGCGACATATGACGCGTCGGTGTGACTTTCTGTGTCTGTGTCA 2880

Qy  2881 CGTAGCTCCTCTATGTTAACATCCAGTGCAGCATCGTTCTTTGGGACTGACTGACTCTAGGGT 2940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881 CGTAGCTCCTCTATGTTAACATCCAGTGCAGCATCGTTCTTTGGGACTGACTGACTCTAGGGT 2940

Qy  2941 TAGGGCCCATAGAAGTGTTGGTGCGTCTCCGAAAGTGTCATGCTTGGTGTATTCTCATTATGT 3000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 TAGGGCCCATAGAAGTGTTGGTGCGTCTCCGAAAGTGTCATGCTTGGTGTATTCTCATTATGT 3000

Qy  3001 TCTTAAGTTTTCTCCTCTAGTGTTTGGTGAGAATGGTGTGTTTTTCTATAAGCACTTGCA 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 TCTTAAGTTTTCTCCTCTAGTGTTTGGTGAGAATGGTGTGTTTTTCTATAAGCACTTGCA 3060

Qy  3061 TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCATTTTTCCC 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCATTTTTCCC 3120

Qy  3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACGCTTGGCGTGTGGGGACAC 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACGCTTGGCGTGTGGGGACAC 3180

Qy  3181 GGTGATGGTTGCCCGTGTGCGCTCGGCGACCTTGTTTTCGCAGGGGTTGGCTAT 3240
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 GGTGATGGTTGCCCGTGTGCGCTCGGCGACCTTGTTTTCGCAGGGGTTAGCTAT 3240

Qy  3241 GCCGCCAGATGGGGTGGGCCATTACCGCACCTTTTACGCTGCAGTGTCTCTCTGAACGTGG 3300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 GCCGCCAGATGGGGTGGGCCATTACCGCACCTTTTACGCTGCAGTGTCTCTCTGAACGTGG 3300

Qy  3301 CACGCTGTCAGCGATGGCAGTGCTCATGACTGGTATAGACCCCGAACTTGGACTTGGAAC 3360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 CACGCTGTCAGCGATGGCAGTGCTCATGACTGGTATAGACCCCGAACTTGGACTTGGAAC 3360

Qy  3361 TATCTTCAGATTAGGATCTCTGGCCACTAGCTACTACATGTCATTGTTTGTGACAACGTGTT 3420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 TATCTTCAGATTAGGATCTCTGGCCACTAGCTACTACATGTCATTGTTTGTGACAACGTGTT 3420

Qy  3421 GTATACTGCTCACCATGGCAGCAGGGGCCGCAAGGGGCGCCGGTTGGCTCATCCCACAGGCTCTATACA 3480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421 GTATACTGCTCACCATGGCAGCAGGGGCCGCAAGGGGCGCCGGTTGGCTCATCCCACAGGCTCCATACA 3480

Qy  3481 CCCAATAACCGTTGACGCGGCTAATGACGCGGCTAATGACCAGGACATCATCAACCACCATGTGGAGCTGG 3540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481 CCCAATAACCGTTGACGCGGCTAATGACGCGGCTAATGACCAGGACATCATCAACCACCATGTGGAGCTGG 3540

Qy  3541 GTCCCTTACTCGGTGCTCTTGCGGGAGACCAAGGGGTATCGGGGTAACACGACTGGGGTC 3600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541 GTCCCTTACTCGGTGCTCTTGCGGGAGACCAAGGGGTATCGGGGTAACACGACTGGGGTC 3600

Qy  3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGCGGGGGCCCTTCCCAT 3660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGCGGGGGCCCTTCCCAT 3660

Qy  3661 GGCTGTTGCCAAGGGTTCTTCAGGGTGCCCCGATTCGTGCGCCCTGCTCGTGCTCTCCGGAGCCATGTTATTGG 3720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661 GGCTGTTGCCAAGGGTTCTTCAGGGTGCCCCGATTCGTGCGCCCTGCTCGTGCTCTCCGGAGCCATGTTATTGG 3720

Qy  3721 GATGTTCACCGCTGGCTGGGATGAAATTCTGCCGGGCCGTTCAGTCAGTTCAGATTAGGGTTAGGCCGTT 3780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721 GATGTTCACCGCTGGCTGGGATGAAATTCTGCCGGGCCGTTCAGTCAGTTCAGATTAGGGTTAGGCCGTT 3780

Qy  3781 GGTGTGTGCTGGGATACCATCCCCAGTACGACAGACATGCCACTCTTGATACAAACCTAC 3840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781 GGTGTGTGCTGGGATACCATCCCCAGTACGACAGACATGCCACTCTTGATACAAACCTAC
```

```
Qy   4921 TTATGTTTTTGTTGACTGAGCCCAACTACACAACTGTGTCATCAGTATGGCTATGCTGCTCC 4980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 TTATGTTTTTGTTGACTGAGCCCAACTACACAACTGTGTCATCAGTATGGCTATGCTGCTCC 4980

Qy   4981 CAATGACGCACCACGGTGCAGGGAGCCCGGCGTTGGGAAAAACCTTGTGGGGGTTCTGTG 5040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 CAATGACGCACCACGGTGCAGGGAGCCCGGCGTTGGGAAAAACCTTGTGGGGGTTCTGTG 5040

Qy   5041 GCGCTTGGACGGCGCTGACGCGTCTCGGCCCAGAGCCCAGGCGAGGTGACCAGATACCA 5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041 GCGCTTGGACGGCGCTGACGCGTCTCGGCCCAGAGCCCAGGCGAGGTGACCAGATACCA 5100

Qy   5101 AATGTGCTTCACTGAAGTCAATACTTCTGGGACAGCCGCACTCGCTGTTGGCGTTGGAGT 5160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101 AATGTGCTTCACTGAAGTCAATACTTCTGGGACAGCCGCACTCGCTGTTGGCGTTGGAGT 5160

Qy   5161 GGCTATGCTTATCTAGCCATTGACACTTTGGCGCCACTTTGTGCGCGGTGCTGCTGGTC 5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161 GGCTATGCTTATCTAGCCATTGACACTTTGGCGCCACTTTGTGCGCGGTGCTGCTGGTC 5220

Qy   5221 TATTACATCAGTCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAAGAAATCGT 5280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221 TATTACATCAGTCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAAGAAATCGT 5280

Qy   5281 GGAGGAGTGTGCATCATTCATTCCCTGGAGGCCATGGTTGCTGCAATTGACAAGCTGAA 5340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5281 GGAGGAGTGTGCATCATTCATTCCCTGGAGGCCATGGTTGCTGCAATTGACAAGCTGAA 5340

Qy   5341 GAGTACAATCACCACAACTAGTCCTTTCACATTGGAAACCGCCCTTGAAAAACTTAACAC 5400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5341 GAGTACAATCACCACAACTAGTCCTTTCACATTGGAAACCGCCCTTGAAAAACTTAACAC 5400

Qy   5401 CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCGCTGTGTGTGTTGTTACTACT 5460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5401 CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCGCTGTGTGTGTTGTTACTACT 5460

Qy   5461 CACTTTACCTGACAATCCCTTGTGCTTTCATTGCGGGTATTACTAC 5520
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   5461 CACTTTACCTGACAATCCCTTGTGCTTTCATTGCGGGTATTACTAC 5520

Qy   5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTTCATTATTGGAGGCGCAATTGGCGTCCAA 5580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTTCATTATTGGAGGCGCAATTGGCGTCCAA 5580

Qy   5581 GCTTACAGACGCTAGAGGCGCACTGGCGTTCATGATGGCCGGGAACAGCTCT 5640
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5581 GCTTACAGACGCTAGAGGCGCACTGGCGTTCATGATGGCCGGGAACAGCTCT 5640

Qy   5641 TGGTACATGGACATCGGTGGGTTTGCTTTGACAAGTGCATGCTCGTGCCGCCTC 5700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5641 TGGTACATGGACATCGGTGGGTTTGTCTTTGACAAGTGCATGCTCGTGCCGCCTC 5700

Qy   5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGGCCCACTATGGATCAGCT 5760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGGCCCACTATGGATCAGCT 5760

Qy   5761 TGCTGGTTTAGTCTACTCCGGTTCAATCCGGCGCCAGAGTTGTGGGCGTCTTGTCAGC 5820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5761 TGCTGGTTTAGTCTACTCCGGTTCAATCCGGCGCCAGAGTTGTGGGCGTCTTGTCAGC 5820

Qy   5821 TTGTGCAATGTTGCTTTGACAACAGCAGGGCCAGATCGACTACTGCGCCAACAGACTTCTTAC 5880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5821 TTGTGCAATGTTGCTTTGACAACAGCAGGGCCAGATCGACTACTGCGCCAACAGACTTCTTAC 5880

Qy   5881 TATGCTTGCTAGGAGCAACAACTGTGTATGAGTCACTGTTTATTGCCACTGTGACATCCG 5940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5881 TATGCTTGCTAGGAGCAACAACTGTGTATGAGTCACTGTTTATTGCCACTGTGACATCCG 5940

Qy   5941 CAGGAAGATACTGGGCATTCTGGAGGCATTCTACCCCCTGGAGTGTCATATCAGCTTGCAT 6000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5941 CAGGAAGATACTGGGCATTCTGGAGGCATTCTACCCCCTGGAGTGTCATATCAGCTTGCAT 6000
```

```
Db   3781 GGTGTGTGCTGGATACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAAACCTAC 3840

Qy   3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGGCAGCGGCAAGTCAAC 3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 TGTGCCTAACGAGTATTCAGTGCAAATTTTAATTGCCCCCACTGGGCAGCGGCAAGTCAAC 3900

Qy   3901 CAAATTACCACTTTCTTACATGTCAGGAGAAGTATGAGGTCTTGGTTCCTAAATCCCAGTGT 3960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 CAAATTACCACTTTCTTACATGTCAGGAGAAGTATGAGGTCTTGGTTCCTAAATCCCAGTGT 3960

Qy   3961 GGCTACAACAGCATCAATGCCAAAGTACATGCACGGCGACGTACGGCGTGAATCAAATTG 4020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 GGCTACAACAGCATCAATGCCAAAGTACATGCACGGCGACGTACGGCGTGAATCAAATTG 4020

Qy   4021 CTATTTTAATGCAAATGTACCAACACAGAGGGCTTCACTTACGTACGACACATATGGCAT 4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 CTATTTTAATGCAAATGTACCAACACAGAGGGCTTCACTTACGTACGACACATATGGCAT 4080

Qy   4081 GTACCTGACCCGGAGCATGTTCCCGGAACTATGATGTGAATCATTTGTGACGAATGCCATGC 4140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 GTACCTGACCCGGAGCATGTTCCCGGAACTATGATGTGAATCATTTGTGACGAATGCCATGC 4140

Qy   4141 TACCGGTGCAACCACCGTGCTTGGCGCCACGGTACCCCCCGGAGTAATCCTCCATCCACCACATGC 4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 TACCGGTGCAACCACCGTGCTTGGCGCCACGGTACCCCCCGGAGTAATCCTCCATCCACCACATGC 4200

Qy   4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCTGGAGTAATCCTCCTTACTACACACCAGTGCA 4260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCTGGAGTAATCCTCCTTACTACACACCAGTGCA 4260

Qy   4261 CAACATAACTGAGATTCAATAACCGATGAAGGCACTATCCCTTTCATGTGAAAAGAGAT 4320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 CAACATAACTGAGATTCAATAACCGATGAAGGCACTATCCCTTTCATGTGAAAAAGAAT 4320

Qy   4321 TAAGGAGGAAAATCTGAAGAAGGGAGACACACCTTATCTTTGAGGCTACCAAAAAACACTG 4380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321 TAAGGAGGAAAATCTGAAGAAGGGAGACACACCTTATCTTTGAGGCTACCAAAAAACACTG 4380

Qy   4381 TGATGAGCTTGCTACGAGTTAGCTCGAAGGAATAACAGCTGTCTCTTACTATAGGGG 4440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 TGATGAGCTTGCTACGAGTTAGCTCGAAGGAATAACAGCTGTCTCTTACTATAGGGG 4440

Qy   4441 ATGTGACATCTCAAAAATCCCTGAGGCGCACTGTGTAGTTGCCACTGGCGTTCGTGTG 4500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 ATGTGACATCTCAAAAATCCCTGAGGCGCACTGTGTAGTTGCCACTGGCGTTCGTGTG 4500

Qy   4501 TACAGGGTACACTGGTGACCTTTGATTCCGTATGACTGTCAGCCCATGTGTGTGTAGAGGCAC 4560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 TACAGGGTACACTGGTGACCTTTGATTCCGTATGACTGTCAGCCCATGTGTGTGTAGAGGCAC 4560

Qy   4561 ATGCCATGTGTGACCCTACTTTCACCATGGCGTGTCGGTGTTCGTGTGGCGGGGTCTTCAGC 4620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 ATGCCATGTGTGACCCTACTTTCACCATGGCGTGTCGGTGTTCGTGTGGCGGGGTCTTCAGC 4620

Qy   4621 AATAGTTAAAGGCCAGCCGTAGGGCCGCACAGGCCGCACAGGCCGCATATACTACTA 4680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4621 AATAGTTAAAGGCCAGCCGTAGGGCCGCACAGGCCGCACAGGCCGCATATACTACTA 4680

Qy   4681 TGTAGACGGGAGTTGTACCCCCTCGGGTATGTTCCTGAATGCAACATTGTTGAAGCCTT 4740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4681 TGTAGACGGGAGTTGTACCCCCTCGGGTATGTTCCTGAATGCAACATTGTTGAAGCCTT 4740

Qy   4741 CGACGCAGCCAAGGCCAAACCCAACCTGGGTTGTTGTCATCAACAGAGCTCAAACTATTCTGACAC 4800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 CGACGCAGCCAAGGCCAAACCCAACCTGGGTTGTTGTCATCAACAGAGCTCAAACTATTCTGACAC 4800

Qy   4801 CTATCGCACCCAACCTGGGTTACCTTGCGATAGGAGCAACACTGTATGACGAGTGGCTGATCT 4860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 CTATCGCACCCAACCTGGGTTACCTTGCGATAGGAGCAACACTGTATGACGAGTGGCTGATCT 4860

Qy   4861 CTTTTCTATGGTCAACCCGAACCTTCATTTGTCAATACTGCAAAAAGACTGCTGACAA 4920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861 CTTTTCTATGGTCAACCCGAACCTTCATTTGTCAATACTGCAAAAAGACTGCTGACAA 4920
```

```
Qy  7081 TTTACCCAGTTACCCTCCAAAAAGGAGGTCTCGAAATGGTCAGACGGAAGTTGGTCAAC 7140
         |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db  7081 GGCTACACACCGCTTCCAGCTACGTTACTGCCCCCGTACCCTAAGATACGGGGAAAAGGA 7140

Qy  7141 GGCTACACACCGCTTCCAGCTACGTTACTGCCCCCGTACCCTAAGATACGGGGAAAAGGA 7200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141 GACTACACAACCGCTTCCAGCTACGTTACTGCCCCCGTACCCTAAGATACGGGGAAAAGGA 7200

Qy  7201 TTCCACTCAGTCAGCCCCGCCAAACGGCCTACAAAAAGAAGTGGGAAAGAGAGTGAGTT 7260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201 TTCCACTCAGTCAGCCCCGCCAAACGGCCTACAAAAAGAAGTGGGAAAGAGAGTGAGTT 7260

Qy  7261 TTCGTGCGACATGGCTACACCTGGACCCACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261 TTCGTGCGACATGAGCTACACTTGGACCCACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320

Qy  7321 TCTGTCTGCAACTCGGGCCATCACTAGTCGTTCCTCAAACAAAGATCATTGGTGTATGT 7380
         |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db  7321 TCTGTCTGCAACTCGGGCCATCACTAGTCGTTCCTCAAACAACAAAGATCATTGGTATGT 7380

Qy  7381 GACTGAGCCGCGGAGCTTAGAAACAAAAAGTCACTATTAATAGACAACCTCT 7440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381 GACTGAGCCGCGGAGCTTAGAAACAAAAAGTCACTATTAATAGACAACCTCT 7440

Qy  7441 GTTCCCCCATCATACCACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441 GTTCCCCCATCATACCACAAGCAAGTGAGATTGGCTAAGGAAAAAAGCTTCAAAAGTTGT 7500

Qy  7501 CGGTGTCATGTGGGACTATGAAGCAGCTGAAGAGTAGCAGCTCACACGCCCTCTAAGTC 7560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7501 CGGTGTCATGTGGGACTATGAAGCAGCTGAAGAGTAGCAGCTCACACGCCCTCTAAGTC 7560

Qy  7561 CCACATCACTGGCCTTCGGGGCACTGATGTTCGTTCTGGGGCACTCGGGGCTGTTCT 7620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7561 CCACATCACTGGCCTTCGGGGCACTGATGTTCGTTCTGGGGCACTCGGGGCTGTTCT 7620

Qy  7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTCATTATCGGCAAACTGTGAT 7680
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTCATTATCGGCAAACTGTGAT 7680

Qy  7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAAACAACAAAGAAACCCCCAAG 7740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAAACAACAAAGAAACCCCCAAG 7740

Qy  7741 GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGT 7800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7741 GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGT 7800

Qy  7801 TGCTCCCTGACGTAGTTAAAGCTGTTCATGGGAGTGCGTACGGGTTTGTGACCCACGTAC 7860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7801 TGCTCCCTGACGTAGTTAAAGCTGTTCATGGGAGTGCGTACGGGTTTGTGACCCACGTAC 7860

Qy  7861 CCGTGTCAAGGCGTCTGTTGTCGATGGCGTCACCCGAGCCACATGCGCCGCGATAC 7920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7861 CCGTGTCAAGGCGTCTGTTGTCGATGGCGTCACCCGAGCCACATGCGCCGCGATAC 7920

Qy  7921 AGTGTGTTTGACAGTACCAGCATCACACCCGAGGATATCATGTGGAGACAGACATCTACTC 7980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7921 AGTGTGTTTGACAGTACCAGCATCACACCCGAGGATATCATGTGGAGACAGACATCTACTC 7980

Qy  7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGGAGCATTGCACACCGAGGGCAGTTATA 8040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGGAGCATTGCACACCGAGGGCAGTTATA 8040

Qy  8041 CGCTTGGAGGACCGATGATCCGTTATGATGGCCGAGATCGGATATCGTAGGTGTAGGTC 8100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8041 CGCTTGGAGGACCGATGATCCGTTATGATGGCCGAGATCGGATATCGTAGGTGTAGGTC 8100

Qy  8101 TTTCCGCGTCTATACTACCTCAAGTTGACCTGCTGGCTCGGCTGGAAGGTAAATGC 8160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8101 TTTCCGCGTCTATACTACCTCAAGTTTGACCTGCTGGCTCGGCTGGAAGGTAAATGC 8160

Qy  8161 TGCAGCGAACAGGTGGCATGGAGAAGAACCCCTCGCTTCCTTATTTGCGGGCATGGATTGTCGAC 8220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161 ...
```

```
Qy  6001 CCGTTGGCTCCACACCCGACGGAGGATGATTGCGGCCTCATTGCTTGCGGGTCTCTAGAGAT 6060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001 CCGTTGGCTCCACACCCGACGGAGGATGATTGCGGCCTCATTGCTTGCGGGTCTCTAGAGAT 6060

Qy  6061 TTGGCAGTATGTGTGCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGCTGGAGTTCA 6120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061 TTGGCAGTATGTGTGCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGCTCGAGTTCA 6120

Qy  6121 GAGCATGGTTAACATTCCTGGTTGTCCTTCTACAGCTGCCGAAGAGGGCTACAAGGGCCC 6180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121 GAGCATGGTTAACATTCCTGGTTGTCCTTCTACAGCTGCCGAAGAGGGCTACAAGGGCCC 6180

Qy  6181 CTGGATTGGATCAGGTATGCTCCAAGCACGTGTCCATGCGGTGCTGAACTCATCTTTTC 6240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181 CTGGATTGGATCAGGTATGCTCCAAGCACGTGTCCATGCGGTGCTGAACTCATCTTTTC 6240

Qy  6241 TGTTGAGAATGGTTTTGCAAAACTTTACAAAGGACCCAGAAACTTGTTCAAATTACTGGAG 6300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241 TGTTGAGAATGGTTTTGCAAAACTTTACAAAGGACCCAGAAACTTGTTCAAATTACTGGAG 6300

Qy  6301 AGGGGCTGTTCCAGTCAACGCTAGGCTGTCGTGTGGGTCGCTAGACCGGACCCAACTGATTG 6360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301 AGGGGCTGTTCCAGTCAACGCTAGGCTGTCGTGTGGGTCGCTAGACCGGACCCAACTGATTG 6360

Qy  6361 GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAAATATGAGAAAATGGGAGA 6420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361 GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAAATATGAGAAAATTGGGAGA 6420

Qy  6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTCACCCAGGTGCCCCCAAC 6480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTCACCCAGGTGCCCCCAAC 6480

Qy  6481 CTTGAGAGCTGCAGTGGCCGTGACGACGGCGTACAGGTTTATCTAGGTGAGCCCCAA 6540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481 CTTGAGAGCTGCAGTGGCCGTGACGACGGCGTACAGGTTTATCTAGGTGAGCCCCAA 6540

Qy  6541 AACTCCTTGGACACCATCTGCTTGTCTGTTACGCGTCCTGCGAGCGGTAAGGTAAAACTGTTAA 6600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541 AACTCCTTGGACACCATCTGCTTGTCTGTTACGCGTCCTGCGAGCGGTAAGGTAAAACTGTTAA 6600

Qy  6601 GCTTCCCTTCCGGTTGACGGTCACACACCTGGTGTGCGGCATGCAACTTAATTTGCGTGA 6660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601 GCTTCCCTTCCGGTTGACGGTCACACACCTGGTGTGCGGCATGCAACTTAATTTGCGTGA 6660

Qy  6661 TGCACTTGAGACAAATGACTGTAATTCCACAAACAACAACTCCTAGTGATGAAGCCGCAGT 6720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661 TGCACTTGAGACAAATGACTGTAATTCCACAAACAACAACTCCTAGTGATGAAGCCGCAGT 6720

Qy  6721 GTCCGCTCTGTTTTCAAACGGAGTTGCGCGGTACAAACCAATTGCTTGAGGCAATTTC 6780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721 GTCCGCTCTGTTTTCAAACGGAGTTGCGCGGTACAAACCAATTGCTTGAGGCAATTTC 6780

Qy  6781 AGTGGGCGTTGACAACCACCAAACTGCCAGCCCCCTCCATCGAAGAGGTAGTGTAAGAAA 6840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781 AGTGGGCGTTGACAACCACCAAACTGCCAGCCCCCTCCATCGAAGAGGTAGTGTAAGAAA 6840

Qy  6841 GCGCCAGTTCCGGGCAAGAACTGTTCGCTTACCTTGCCTCCCCCCCGAGATCCGTCCC 6900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841 GCGCCAGTTCCGGGCAAGAACTGTTCGCTTACCTTGCCTCCCCCCCGAGATCCGTCCC 6900

Qy  6901 AGGAGTGTCATGTCCTGAAAGCCTGACAACGAGTTAGAAGGTCCTTCAAACCT 6960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901 AGGAGTGTCATGTCCTGAAAGCCTGACAACGAGTTAGAAGGTCCTTCAAACCT 6960

Qy  6961 CCCTCCTTCCACCACCTGTTCTACAGTTGGCCATGCCGATGCCGGGGAGCCCGGGTGA 7020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961 CCCTTCCTTCCACCACCTGTTCTACAGTTGGCCATGCCGATGCCGGGGAGCCCGGGTGA 7020

Qy  7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGGAAACAGGCGGCCCTGATGA 7080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGGAAACAGGCGGCCCTGATGA 7080

Qy  7081 TTTACCCAGTTACCCTCCAAAAAGGAGGTCTCGAATGGTCAGACGAAAGTTGGTCGAC 7140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081 TTTACCCAGTTACCCTCCAAAAAGGAGGTCTCGAATGGTCAGACGAAAGTTGGTCGAC 7140
```

KEYWORDS
SOURCE　　　　Unknown.
  ORGANISM　　Unknown.
　　　　　　　Unclassified.
REFERENCE　　1　(bases 1 to 9143)
  AUTHORS　　 Simons,J.N., Pilot-Matias,T.J., Dawson,G.J., Schlauder,G.G.,
　　　　　　　Desai,S.M., Leary,T.P., Muerhoff,A.S., Erker,J.C., Buijk,S.L. and
　　　　　　　Mushahwar,I.K.
  TITLE　　　 Non-A, non-B, non-C, non-D, non-E hepatitis reagents and methods
　　　　　　　for their use
  JOURNAL　　 Patent: US 6558898-A 393 06-MAY-2003;
FEATURES　　　　　　　 Location/Qualifiers
  source　　　　　　　 1..9143
　　　　　　　　　　　 /organism="unknown"
　　　　　　　　　　　 /mol_type="genomic DNA"
ORIGIN

Query Match　　96.4%;　Score 9059.8;　DB 6;　Length 9143;
Best Local Similarity　99.6%;　Pred. No. 0;
Matches 9096;　Conservative　0;　Mismatches　37;　Indels　4;　Gaps　1;

```
Qy      1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCTGGAGCACCCCCCTAG     60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCTGGAGCACCCCCCTAG     60

Qy     61 CAGGGCGTGGGGGATTCCCCTGCCCGTCTGCGAGAAGGGTGGAGCCAACCACCTTAGTAT    120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAGGGCGTGGGGGATTCCCCTGCCCGTCTGCGAGAAGGGTGGAGCCAACCACCTTAGTAT    120

Qy    121 GTAGGCGGCGGGACTCATGAGCGTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGC    180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTAGGCGGCGGGACTCATGAGCGTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGC    180

Qy    181 CCTGATGGCGGCGTTCATGGGTTCGGTGGTGGTGGCGGCTTAGGCAGCCTCCAGCGCCACCA    240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCTGATGGCGGCGTTCATGGGTTCGGTGGTGGTGGCGGCTTAGGCAGCCTCCAGCGCCACCA    240

Qy    241 CCTCCCAGATAGAGCGGCACTGTAGGGAGCGAGAGACCGGGACCGGTCACTACTACCAAGGACG    300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTCCCAGATAGAGCGGCACTGTAGGGAGCGAGAGACCGGGACCGGTCACTACTACCAAGGACG    300

Qy    301 CAGACCCTCTTTTGAGTATCACGCCTCCGGAGTAGTTGGGCAAGCCCACCTATATGTGT    360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGACCCTCTTTTGAGTATCACGCCTCCGGAGTAGTTGGGCAAGCCCACCTATATGTGT    360

Qy    361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT    420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT    420

Qy    421 CTGGGAGTCTCGTAGACCGTAGCACATGCACGTGTATTCTACTCAAACAAGTCCTGTACC    480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGGGAGTCTCGTAGACCGTAGCACATGCACGTGTATTCTACTCAAACAAGTCCTGTACC    480

Qy    481 TGCGCCCAGAACGCGCAAGAACAGCAGACGCAGACAGCTTCATATCCTGTGTCCATTAAAAC    540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGCGCCCAGAACGCGCAAGAACAGCAGACGCAGACAGCTTCATATCCTGTGTCCATTAAAAC    540

Qy    541 ATCTGTTGAAAGGGGACAGCCAAAGTCCAGCGCGATGCGTCGGCCTCGTAA    600
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ATCTGTTGAAAGGGGACAGCCAAAGTCCAGCGCGATGCGTCGGCCTCGTAA    600

Qy    601 TTACAAAATTGCTGGTATCCATGATGGCTTCAGCGACATTGGCTTCAGCGC    660
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTACAAAATTGCTGGTATCCATGATGGCTTCAGCGACATTGGCTTCAGCGC    660

Qy    661 TCATGGTTGGGACGCCAAGACCCTCGCCATAAGTCTCGCAATCTTCTGA    720
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TCATGGTTGGGACGCCAAGACCCTCGCCATAAGTCTCGCAATCTTCTGA    720

Qy    721 TTACCCTTGGGGTGGGATTGGTGATGTTACAACTCACACCTCCAGTAGGCCCGCTGGT    780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TTACCCTTGGGGTGGGATTGGTGATGTTACAACTCACACCTCCAGTAGGCCCGCTGGT    780
```

```
Db   8161 TGCAGCCGACAGGCTGGCATGAAGAACCCTCGCTTCCTTATTTGCGGCGATGATTGCAC    8220
Qy   8221 CGTAATTTGGAAGAGCGCCGGAGCAGATGCAGCTGCAGCAGCTGGTGTCTTTGCTAG    8280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8221 CGTAATTTGGAAGAGCGCCGGAGCAGATGCAGCTGCAGCAGCTGGTGTCTTTGCTAG    8280

Qy   8281 CTGGATGAAGTGATGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACAGTTTGGA    8340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8281 CTGGATGAAGTGATGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACAGTTTGGA    8340

Qy   8341 AGAATTAACATCATGCTCATCAAATGTTAACTTCTGAATTACCAAAAGTGGCAAGCCTTA    8400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8341 AGAATTAACATCATGCTCATCAAATGTTAACTTCTGAATTACCAAAAGTGGCAAGCCTTA    8400

Qy   8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTGCCGAGGGTCTGGG    8460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTGCCGAGGGTCTGGG    8460

Qy   8461 ATACAACCCCAGTGCGTGCGTGTCATCTAATACATCACAGCTAGGCCATGTTTGTGGGT    8520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8461 ATACAACCCCAGTGCGTGCGTGTCATCTAATACATCACAGCTAGGCCATGTTTGTGGGT    8520

Qy   8521 TAGCCGTGTGTGTTGGCTGTCCATTTCATGCAGCAGATGCTCTTTGAGGACAAAACTTCCCGA    8580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8521 TAGCCGTGTGTGTTGGCTGTCCATTTCATGCAGCAGATGCTCTTTGAGGACAAAACTTCCCGA    8580

Qy   8581 GACTGTGACCTTGACTGGTATGGGAAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG    8640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8581 GACTGTGACCTTGACTGGTATGGGAAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG    8640

Qy   8641 CATCATTGCTGGTGTGCACGGTATTGAGGCTTTCTCGGTGCGTGTGCCTACACCAACGCTGA    8700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8641 CATCATTGCTGGTGTGCACGGTATTGAGGCTTTCTCGGTGCGTGTGCCTACACCAACGCTGA    8700

Qy   8701 GATCCTCAGAGTTCCCAATCACTACAGACAGATGACCATGCCATGCATGCCTGCGGCCGC    8760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8701 GATCCTCAGAGTTCCCAATCACTACAGACAGATGACCATGCCATGCATGCCTGCGGCCGC    8760

Qy   8761 AAAGAAAGCCAGGGCCGTCCTGCCCAGCGCCAAGAGCCGTGCGGAGCACACGCAAATT    8820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8761 AAAGAAAGCCAGGGCCGTCCTGCCCAGCGCCAAGAGCCGTGCGGAGCACACGCAAAATT    8820

Qy   8821 GGCTCGCTTCCTTCTCTGGCATGCTACATCTAGACCCTACCAGATTTGGATAAGACGAG    8880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8821 GGCTCGCTTCCTTCTCTGGCATGCTACATCTAGACCCTACCAGATTTGGATAAGACGAG    8880

Qy   8881 CGTGGGCTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGGAGGGGATGTGTT    8940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8881 CGTGGGCTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGGAGGGGATGTGTT    8940

Qy   8941 TATTACCACCAGAGAGAGATTGCAGAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTTT    9000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8941 TGTTACCACCAGAGAGAGAAGAGTTGCAGAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTTT    9000

Qy   9001 TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA    9060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9001 TGCCCTAGGGCTCATTGCTGTTGGACTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA    9060

Qy   9061 CTAACAG-----TTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC    9116
          |||||||     ||||||||||||||||||||||||||||||||||||||||||||||||
Db   9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC    9120

Qy   9117 GGGCTTAACGACCCCGC    9133
          |||||||||||||||||
Db   9121 GGGCTTAACGACCCCGC    9137
```

RESULT 11
AR310129
LOCUS　　　　　AR310129　　　　　　9143 bp　　　DNA　　　　　linear　　　PAT　12-JUN-2003
DEFINITION　　Sequence 393 from patent US 6558898.
ACCESSION　　 AR310129
VERSION　　　　AR310129.1　GI:31702407

```
Qy   781  GGCAGGAGGCGGTCGTTCGACCAGTCTGCCAGATAGTACGCTTCGCTGGAGGATGGAGTCAA  840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  GGCAGGAGGCGGTCGTTCGACCAGTCTGCCAGATAGTACGCTTCGCTGGAGGATGGAGTCAA  840

Qy   841  CTGGGCTACTGGTGTCGGAGAGTGCTTCGGTGTCCACCTTTTGTGTGTATGTCTGCTATCTTTGGGCCTG  900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  CTGGGCTACTGGTGTCGGAGAGTGCTTCGGTGTCCACCTTTTGTGTGTATGTCTGCTATCTTTGGGCCTG  900

Qy   901  TCCCTGTAGTGGGGCGGCGGGGTCACTGACCTGCACTGACCAGGACACAAATACCACAATTG  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  TCCCTGTAGTGGGGCGGCGGGGTCACTGACCTGCACTGACCAGGACACAAATACCACAATTG  960

Qy   961  CTGCCAGCGTAATCAGTTATCTATTGTTCTCCTTGCCACTTGCCTACACGAGCCTGGTTG  1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  CTGCCAGCGTAATCAGTTATCTATTGTTCTCCTTGCCACTTGCCTACACGAGCCTGGTTG  1020

Qy  1021  TGTGATCTGTGCGGACGAGTGCTGTGGGTTCCGCCAATCGTACATCTTCACACCCTTCCAA  1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  TGTGATCTGTGCGGACGAGTGCTGTGGGTTCCGCCAATCGTACATCTTCACACCCTTCCAA  1080

Qy  1081  TTTGGACTGGCACGGACGACTCCTTCTGGCTGACCACCATTGATTTTGTTATGGGGGCTCTTGT  1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  TTTGGACTGGCACGGACGACTCCTTCTGGCTGACCACCATTGATTTTGTTATGGGGGCTCTTGT  1140

Qy  1141  GACCTGTGACGCGCCCTTGACATTGGTGAGTTGCGTGTGTGTATTAGTCGGTGACTG  1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  GACCTGTGACGCGCCCTTGACATTGGTGAGTTGCGTGTGTGTATTAGTCGGTGACTG  1200

Qy  1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTGTTACCTGA  1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTGTTACCTGA  1260

Qy  1261  AGTGCCCACTGGAATAGATCCTGGGTTTATCGGATGGTTCCTAGGGTTATCGGATGGCCGGCAGGT  1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  AGTGCCCACTGGAATAGATCCTGGGTTTATCGGATGGTTCCTAGGGTTATCGGATGGCCGGCAGGT  1320

Qy  1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAGCTTCCATAGACGCTATTGCGACTAT  1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAGCTTCCATAGACGCTATTGCGACTAT  1380

Qy  1381  GTTTAGCAGTGTACACTACCTGCGGTTGGCGCTCTGATCTACTATGCCTCTCGGGGCAA  1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  GTTTAGCAGTGTACACTACCTGCGGTTGGCGCTCTGATCTACTATGCCTCTCGGGGCAA  1440

Qy  1441  GTGGTATCAGTTGCTCCTAGCGCTTATGCTTTACATGAAGAAGCGACCTTCTGGAAACCCCAT  1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  GTGGTATCAGTTGCTCCTAGCGCTTATGCTTTACATGAAGAAGCGACCTTCTGGAAACCCTAT  1500

Qy  1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCTTGATGATACCATGTCC  1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCTTGATGATACCATGTCC  1560

Qy  1561  TTTGCCACTCTTATTTGAGTGAGAATGTCATTTGTTACAGTCCAAAGTGGAC  1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  TTTGCCACTCTTATTTGAGTGAGAATGTCATTTGTTACAGTCCAAAGTGGAC  1620

Qy  1621  CAGGCCTATCACTAGTAGATATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG  1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  CAGGCCTATCACTAGTAGATATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG  1680

Qy  1681  TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA  1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA  1740

Qy  1741  TGTGCCATCGTACTGCACTGATGCAGTGTGGAACGACACTCGCAACACTTA  1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  TGTGCCATCGTACTGCACTGATGCAGTGTGGAACGACACTCGCAACACTTA  1800

Qy  1801  CGAAGCATGCGGTGTAACACCATGGCTGTAACCGCACAACCGGCTCAGCCCTGAA  1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  CGAAGCATGCGGTGTAACACCATGGCTGTAACCGCACAACCGGCTCAGCCCTGAA  1860

Qy  1861  ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC  1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC  1920

Qy  1921  AGGCCATTGTATTTGAGGGATCAGATCAGATACCCCTATAGTTACTTTTATGACCCTGTGAA  1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  AGGCCATTGTATTTGAGGGATCAGATCAGATACCCCTATAGTTACTTTTATGACCCTGTGAA  1980

Qy  1981  TTCCACTCTCTACCACCGGAGAGGTGGCTAGGTGCCCGGTAGGTTGCCCGGTACCCCACCTGTGGTACG  2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  TTCCACTCTCTACCACCGGAGAGGTGGCTAGGTGCCCGGTAGGTTGCCCGGTACCCCACCTGTGGTACG  2040

Qy  2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG  2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG  2100

Qy  2101  ATTGATCACCAAAGACAAAGCCTGGAAAAATAATCAGTCTTATATTCCGCACCACGGGTGC  2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  ATTGATCACCAAAGACAAAGCCTGGAAAAATAATCAGTCTTATATTCCGCACCACGGGTGC  2160

Qy  2161  TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGTGCTAATTCTGTTGGGGTGTGTGGTGG  2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGTGCTAATTCTGTTGGGGTGTGTGGTGG  2220

Qy  2221  CAGCAAGTATCTTATTTTAGCTACACCTCTGTACTTGTCCCTTGTTTGTTTGGGCGGCCTTC  2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  CAGCAAGTATCTTATTTTAGCTACACCTCTGTACTTGTCCCTTGTTTGTTTGGGCGGCCTTC  2280

Qy  2281  TGGTTACCCTTGCGTCTCCATCCCAGTCGTATCTCCAAGCTGGCTGGGGATGT  2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TGGTTACCCTTGCGTCTCCATCCCAGTCGTATCTCCAAGCTGGCTGGGGATGT  2340

Qy  2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTTGCTTGATTTTCTTCATCGTTGCTATCTCCG  2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTTGCTTGATTTTCTTCATCGTTGCTATCTCCG  2400

Qy  2401  CTGCAGGCTACGTTATGCGTCCCCCTTTAGGGTTTGTGCCCATGCGTCGCGGGCCTTGCCCCT  2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  CTGCAGGCTACGTTATGCGTCCCCCTTTAGGGTTTGTGCCCATGCGTCGCGGGCCTTGCCCCT  2460

Qy  2461  AACTTCTTGTTGCAGCAGCTGTGCCAACCAGATTATGACTGGGGTGGGCTGCTTGTCTTGT  2520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  AACTTCTTGTTGCAGCAGCTGTGCCAACCAGATTATGACTGGGGTGGGCTGCTTGTCTTGT  2520

Qy  2521  AGTGGCAGGTTAGTTTGTGGGGCCGGTAACGGCGTCACCCGCATAGCTCTTGT  2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  AGTGGCAGGTTAGTTTGTGGGGCCGGTAACGGCGTCACCCGCATAGCTCTTGT  2580

Qy  2581  AGGTCCTTGGCCTTCTCGGTAGCGCTTTAACCTCTTGGTTGCATTGGAACGGCGTTCAGC  2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  AGGTCCTTGGCCTTCTCGGTAGCGCTTTAACCTCTTGGTTGCATTGGAACGGCGTTCAGC  2640

Qy  2641  TTTTGATACCGGAGATAATTGGAGGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT  2700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  TTTTGACACCGGAGATAATTGGAGGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT  2700

Qy  2701  GTCTCGTTTTGGCTTCTTTGCTCACTTGTACCTCGCTGCTGTTAGTTAAACTCCTATCT  2760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  GTCTCGTTTTGGCTTCTTTGCTCACTTGTACCTCGCTGCTGTTAGTTAAACTCCTATCT  2760

Qy  2761  TTGCAACGTTGGGAGAATTGGTTTGGAACGTTACACTAAGACCGGAGAGGTTTTTCCT  2820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  TTGCAACGTTGGGAGAATTGGTTTGGAACGTTACACTAAGACCGGAGAGGTTTTCTCCT  2820

Qy  2821  TGTGCGGTTTGTTTCCCCGGTGCGACATATGACGCGCTGGTGACTTTCTGTGTGTCA  2880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  TGTGCGGTTTGTTTCCCCGGTGCGACATATGACGCGCTGGTGACTTTCTGTGTGTCA  2880

Qy  2881  CGTAGCTCTTCTTCTATGTTTAACATCCAGTGCAGCATCGTTCTTGGGACTGACTCTAGGGT  2940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  CGTAGCTCTTCTTCTATGTTTAACATCCAGTGCAGCATCGTTCTTGGGACTGACTCTAGGGT  2940

Qy  2941  TAGGGCCCATAGAATGTTGGCGTCTCCGGAAGTGTCATGCGTATTCTGGTATTCTCATTATGT  3000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  2941 TAGGGCCCATAGAATGTTGGTGCGTCTCGGAAAGTGTCATGCTTGTGTATTCTCATTATGT 3000

Qy  3001 TCTTAAGTTTTCCTCTTAGTGTTTGGTGAGAATGGTGTGTTTTTCTATAAGCACTTGCA 3060
         ||||||||||||||| |||||||||||||||||||||||| ||||||||||||||||||
Db  3001 TCTTAAGTTTTCCCTCTTAGTGTTTGGTGAGAATGGTGTGTTTTTCTATAAGCACTTGCA 3060

Qy  3061 TGGTGATGTCTTGCCTAATGATTTTTGCCTCGAAACTACCATTGCAAGAGCCATTTTTCCC 3120
         ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db  3061 TGGTGATGTCTTGCCTAATGATTTTTGCCTCGAAACTACCATTGCAAGAGCCATTTTTCCC 3120

Qy  3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGAAGACGCTTGGCGTGTGGGGACAC 3180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGAAGACGCTTGGCGTGTGGGGACAC 3180

Qy  3181 GGTTGATGGTTGCCCGTGTTGCGGCTCTCGGCGACCTTGTTTTCGCAGGGTTGCTAT 3240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 GGTTGATGGTTGCCCGTGTTGCGGCTCTCGGCGACCTTGTTTTCGCAGGGTTAGCTAT 3240

Qy  3241 GCCGCCAGATGGTGGGCGCCATTACCGCACCTTTACGCTGCAGTGTCTCTCGACGTGG 3300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 GCCGCCAGATGGTGGGCGCCATTACCGCACCTTTACGCTGCAGTGTCTCTCGACGTGG 3300

Qy  3301 CACGCTGTCAGCGATGGCAGTGGTCATGATGGTATAGACCCCGAACTTGGACTTGGAAC 3360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 CACGCTGTCAGCGATGGCAGTGGTCATGATGGTATAGACCCCGAACTTGGACTGGAAC 3360

Qy  3361 TATCTTCAGATTAGGATCTCTGCCACTAGTACATGGGAATTGTTTGTGACAACGTGTT 3420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 TATCTTCAGATTAGGATCTCTGCCACTAGTACATGGGATTTGTTTGTGACAACGTGTT 3420

Qy  3421 GTATACTGCTCACCATGGCGAGCAAGGGCGCCGGTTGGCTCATCCCACCAGGCTCTATACA 3480
         ||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  3421 GTATACTGCTCACCATGGCGAGCAAGGGCGCCGGTTGGCTCATCCCACCAGGCTCCATACA 3480

Qy  3481 CCCAATAACCGTTGACGCGGCTAATGACCACGACATCTATCAACCACCATGTGGAGCTGG 3540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481 CCCAATAACCGTTGACGCGGCTAATGACCACGACATCTATCAACCACCATGTGGAGCTGG 3540

Qy  3541 GTCCCTTACTCGGTCTCTTCGGTGGAGACCAAGGGTATCTGGTAACAGACTGGGGTC 3600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541 GTCCCTTACTCGGTCTCTTCGGTGGAGACCAAGGGTATCTGGTAACAGACTGGGGTC 3600

Qy  3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGCGGGGCCCTTCCCAT 3660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGCGGGGCCCTTCCCAT 3660

Qy  3661 GGCTGTTGCCAAGGGTTCTTCAGGTTGCCCGGATTCTGCTGTCTGTGCTTCGTGTTATTGG 3720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661 GGCTGTTGCCAAGGGTTCTTCAGGTTGCCCGGATTCTGCTGTCTGTGCTTCGTGTTATTGG 3720

Qy  3721 GATGTTCACCCGCTGCTAGAAATTCTCTGCGGTTCAGTCAGATTAGGGTTAGGCCGTT 3780
         ||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db  3721 GATGTTCACCCGCTGCTAGAAATTCTCTGCGGTTCAGTCAGATTAGGGTTAGGCCGTT 3780

Qy  3781 GGTGTGTGCTGGATACCATCCCCAGTACAGCACATGCCACTCTTGATACAAAACCTAC 3840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781 GGTGTGTGCTGGATACCATCCCCAGTACAGCACATGCCACTCTTTGATACAAAACCTAC 3840

Qy  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCACTGGCCAGCGGCAAGTCAAC 3900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCACTGGCCAGCGGCAAGTCAAC 3900

Qy  3901 CAAATTACCACTTTCTTACATGCAGGAAGAAGTATGAGGTCCTAAATCCCAGTGT 3960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 CAAATTACCACTTTCTTACATGCAGGAAGAAGTATGAGGTCCTAAATCCCAGTGT 3960

Qy  3961 GGCTACAACAGCATCAAATGCCAAAGTACACGCGACTACGGCGTACGGCACCAGATACCA 4020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 GGCTACAACAGCATCAAATGCCAAAGTACACGCGACTACGGCGTACGGCACCAGATACCA 4020

Qy  4021 CTATTTTAATGCAAATGTACCAACACACAGGGGCTTCACTTACGTACAGCACATATGGCAT 4080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 CTATTTTAATGCAAATGTACCAACACACAGGGGCTTCACTTACGTACAGCACATATGGCAT 4080
```

```
Qy  4081 GTACCTGACCGGAGCATGTCCGGAACTATCATGTGTAATCATTGTGACGAATGCCATGC 4140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 GTACCTGACCGGAGCATGTCCGGAACTATCATGTGTAATCATTGTGACGAATGCCATGC 4140

Qy  4141 TACCGATGCAACCACCGTGTGGGCATTGGAAAGGTCCTAACCGAAGTCCATCCAAAAA 4200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 TACCGATGCAACCACCGTGTGGGCATTGGAAAGGTCCTAACCGAAGTCCATCCAAAAA 4200

Qy  4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCTGGAGTAATCCCTACACCACATGC 4260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCTGGAGTAATCCCTACACCACATGC 4260

Qy  4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTACTCCCTTTCATGGAAAAAAGAT 4320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTACTCCCTTTCATGGAAAAAAAGAT 4320

Qy  4321 TAAGGAGGAAAATCTGAAGAAAGGGGACACCTTATCTTTGAGGCTACCAAAAACACTG 4380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321 TAAGGAGGAAAATCTGAAGAAAGGGGACACCTTATCTTTGAGGCTACCAAAAACACTG 4380

Qy  4381 TGATGAGCTTGCTAACGAGTTAGCTGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381 TGATGAGCTTGCTAACGAGTTAGCTGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440

Qy  4441 ATGTGACATCTCAAAAATCCCTGAGGCGACTGTGTAGTAGTGCCACTGATGCCTTGTG 4500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441 ATGTGACATCTCAAAAATCCCTGAGGCGACTGTGTAGTAGTGCCACTGATGCCTTGTG 4500

Qy  4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGGGCTCATGGTAGAAGGCAC 4560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGGGCTCATGGTAGAAGGCAC 4560

Qy  4561 ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTGTGTTCGTGTGTGGGGGGTTTCAGC 4620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561 ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTGTGTTCGTGTGTGGGGGGTTCTCAGC 4620

Qy  4621 AATAGTTAAAGGCCAGCCAGCGTGGGGGAGCGCTGGGGAGCTCGGCATATACTACTA 4680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621 AATAGTTAAAGGCCAGCCAGCGTGGGGGAGCGCTGGGGAGCTCGGCATATACTACTA 4680

Qy  4681 TGTAGACGGAGTTGTACCCCTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGGCCTT 4740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681 TGTAGACGGAGTTGTACCCCTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGGCCTT 4740

Qy  4741 CGACGCAGCCAAGGCATGGTATGGTGTTGTCATCAACAGAAGCTCAAACTATTCTGGACAC 4800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741 CGACGCAGCCAAGGCATGGTATGGTGTTGTCATCAACAGAAGCTCAAACTATTCTGGACAC 4800

Qy  4801 CTATCCGCACCAACCTGGGTTACCTGCGATAGGAGCCAAATTGGACGAGTGGGGCTGATCT 4860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801 CTATCCGCACCAACCTGGGTTACCTGCGATAGGAGCCAAATTGGACGAGTGGGGCTGATCT 4860

Qy  4861 CTTTTCTATGGTCAACCCGAACCTTCATTTGCAAAAGAACTGCTGACAA 4920
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861 CTTTTCTATGGTCAACCCGAACCTTCATTTGCAAAAGAACTGCTGACAA 4920

Qy  4921 TTATGTTTGTTGTGACTGACTGGTATGCTGTCATCAGTATGGCTATGCTGCTCC 4980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921 TTATGTTTGTTGTGACTGACTGGTATGCTGTCATCAGTATGGCTATGCTGCTCC 4980

Qy  4981 CAATGACGCACCACGGTGCCAGGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG 5040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981 CAATGACGCACCACGGTGCCAGGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG 5040

Qy  5041 GCGCTTGGACGCGCTGACGCCTGCCGACGCGCTGGAGGTGACCAGAGTGACCAGATACCA 5100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041 GCGCTTGGACGCGCTGACGCCTGCCGACGCGCTGGAGGTGACCAGAGTGACCAGATACCA 5100

Qy  5101 AATGTGCTTCACTGAAGTCAATACTTCTCGGACAGCAGCCGCTGCGTGTTGTGGCGTTGGAGT 5160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101 AATGTGCTTCACTGAAGTCAATACTTCTCGGACAGCAGCCGCTGCGTGTTGTGGCGTTGGAGT 5160
```

Left column:

```
Qy  5161 GGCTATGGCTTATCTAGCCATTGACACTTTGGCGCCACTTGTGTGCGGTTGCTGTC 5220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161 GGCTATGGCTTATCTAGCCATTGACACTTTGGCGCCACTTGTGTGCGGTTGCTGTC 5220

Qy  5221 TATTACATCAGTCCTACCGGTCCTACCGGTGCTACGTGTCGCCCCAGTGGTTGACGAAGAGAAATCGT 5280
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221 TATTACATCAGTCCTACCGGTCCTACCGGTGCTACGTGTCGCCCCAGTGGTTGACGAAGAGAAATCGT 5280

Qy  5281 GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGGTGCGCAATTGACAAGCTGAA 5340
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281 GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGGTGCGCAATCGATAAGCTGAA 5340

Qy  5341 GAGTACAATCACCACAACTAGTCCTTTCACATTGGAAACCGCCCTTGAAAAACTTAACAC 5400
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341 GAGTACAATAACCACCACTAGTCCTTTCACATTGGAAACCGCCCTTGAAAAACTTAACAC 5400

Qy  5401 CTTTCTTGGGCCTCACTTGTGTCATCATGCCTGCTTCATTGCGGGTATTGCTGTGTGGTTTAGT 5460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401 CTTTCTTGGGCCTCACTTGTGTCATCATGCCTGCTTCATTGCGGGTATTGCTGTGTGGTTAGT 5460

Qy  5461 CACTTTACCTGACAATCCCTTTGCATCATGCGTGTTTTGCTTTCATTGGGGTATTACTAC 5520
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461 CACTTTACCTGACAATCCCTTTGCATCATGCGTGTTTTGCTTTCATTGGGGTATTACTAC 5520

Qy  5521 CCCACTACCTCACAAGATCAAAAATGTTCCTGTCATTATTTGGAGGCGCAATTGCGTCCAA 5580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521 CCCACTACCTCACAAGATCAAAAATGTTCCTGTCATTATTTGGAGGCGCAATTGCGTCCAA 5580

Qy  5581 GCTTACAGACGGTAGAGGCGCACTGGCGTTCATGATGGCCGGGCTGCGGGAACAGTCT 5640
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581 GCTTACAGACGGTAGAGGCGCACTGGCGTTCATGATGGCCGGGCTGCGGGAACAGTCT 5640

Qy  5641 TGGTACTAGGACATCGGTGGGTTTGTCTTGACATGCTAGGCGCTATGCTGCCGCCTC 5700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641 TGGTACTAGGACATCGGTGGGTTTGTCTTGACATGCTAGGCGCTATGCTGCCGCCTC 5700

Qy  5701 ATCCACTGCTTGCTTGACATTTAAAATGCTTGATGGGTGAGTGAGTGGTGAGTGCAGCT 5760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701 ATCCACTGCTTGCTTGACATTTAAAATGCTTGATGGGTGAGTGAGTGGTGAGTGCAGCT 5760

Qy  5761 TGCTGGTTTAGTCTACTCCGCGTTCAATCCGGCCGCCAGGAGTGTGGGCGTCTTGTCAGC 5820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761 TGCTGGTTTAGTCTACTCCGCGTTCAATCCGGCCGCCAGGAGTGTGGGCGTCTTGTCAGC 5820

Qy  5821 TTGTGCAATGTTGCTTGACAACAGCAGGGCGGAGATGATTGCGGCCTCCAACAGACTTCTAC 5880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821 TTGTGCAATGTTGCTTTGACAACAGCAGGGCGGAGATGATTGCGGCCTCCAACAGACTTCTAC 5880

Qy  5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940

Qy  5941 CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT 6000
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941 CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT 6000

Qy  6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGCCTCATTGCTTGGGGTCTAGAGAT 6060
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGCCTCATTGCTTGGGGTCTAGAGAT 6060

Qy  6061 TTGGCAGTATGTGCAATTTCTTGTGATTTGCTTTAATGTCCTTAAAGCTGGAGTTCA 6120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061 TTGGCAGTATGTGCAATTTCTTGTGATTTGCTTTAATGTCCTTAAAGCTGGAGTTCA 6120

Qy  6121 GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC 6180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121 GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAGGGGGTACAAGGGCCC 6180

Qy  6181 CTGGATTGGATCAGGTATGCTCCAAGCACGCTGTTCCATGCGGTGCTGAACTCATCTTTC 6240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181 CTGGATTGGATCAGGTATGCTCCAAGCACGCTGTTCCATGCGGTGCTGAACTCATCTTTC 6240

Qy  6241 TGTTGAGAATGGTTTGCAAAACTTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG 6300
```

Right column:

```
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241 TGTTGAGAATGGTTTGCAAAACTTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG 6300

Qy  6301 AGGGGCTGTTCCAGTCAACGCGTGTGGTCGGCTGTCGGCTGGCTCGACGGCCACCAACTGATTG 6360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301 AGGGGCTGTTCCAGTCAACGCGTGTGGTCGGCTGTCGGCTGGCTCGACGGCCACCAACTGATTG 6360

Qy  6361 GACTAGTCTTGTGTCAATTATGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA 6420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361 GACTAGTCTTGTGTCAATTATGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA 6420

Qy  6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTGTTTCACCCAGGTGCCCCAAC 6480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTGTTTCACCCAGGTGCCCCAAC 6480

Qy  6481 CTTGAGAGCTGCAGTGGCCGTGGACGGGCGGGCGGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481 CTTGAGAGCTGCAGTGGCCGTGGACGGGCGGGCGGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540

Qy  6541 AACTCCTTGGACGACATCTGCTTGCTCCTGACGTTACGGTCCGGACGGGTAAGGGTAAAAACTGTTAA 6600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541 AACTCCTTGGACGACATCTGCTTGCTCCTGACGTTACGGTCCGGACGGGTAAGGGTAAAAACTGTTAA 6600

Qy  6601 GCTTCCCTTCCGGTTGACGGTGCACACACCTGGTGTGGGCATGCAGTCCTAATTTTGCGTGA 6660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601 GCTTCCCTTCCGGTTGACGGTGCACACACCTGGTGTGGGCATGCAGTCCTAATTTTGCGTGA 6660

Qy  6661 TGCACTTGAGACAAATGACTGTAATTCCACAACACACTCCTAGTGATGAAGCCGCAGT 6720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661 TGCACTTGAGACAAATGACTGTAATTCCATAAACACACTCCTAGTGATGAAGCCGCAGT 6720

Qy  6721 GTCCGCTCTTGTGTTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC 6780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721 GTCCGCTCTTGTGTTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC 6780

Qy  6781 AGCTGGCGTTGACACACCAAAACTGCCAGCCCCCATCCAGAAGAGGTAGTGGTAAGAAA 6840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781 AGCTGGCGTTGACACACCAAAACTGCCAGCCCCCATCCAGAAGAGGTAGTGGTAAGAAA 6840

Qy  6841 GCGGCCAGTTCCGGCCAAGAACTCGGTTCGCTTCGCCTTCCCCCTCCCGAGATCCGTCCC 6900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841 GCGGCCAGTTCCGGCCAAGAACTCGGTTCGCTTCGCCTTCCCCCTCCCGAGATCCGTCCC 6900

Qy  6901 AGGAGTGTCATGTCCAAAGCCTGCAACGAAGTGACCCGTTGAGAAGGTCCTTCAAACCT 6960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901 AGGAGTGTCATGTCCAAAGCCTGCAACGAAGTGACCCGTTGAGAAGGTCCTTCAAACCT 6960

Qy  6961 CCCCCCTTCACCCACTGTTCCACAGCTTCTACAGTTGGCCATGCCGATGCCCCGTGTTGGGAGCCGGTGA 7020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961 CCCTTCTTCACCCACTGTTCCACAGCTTCTACAGTTGGCCATGCCGATGCCCCGTGTTGGGAGCCAGGTGA 7020

Qy  7021 GTGTAACCCTTCACTGACAATTGGATGTGCAATTGACGGCGGAGGCGGCCCTGATGA 7080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021 GTGTAACCCTTCACTGACAATTGGATGTGCAATTGACGGCGGAGGCGGCCCTGATGA 7080

Qy  7081 TTTACCCAGTTACCCTCCCAAAAAGGAGGTCTCTGACGTGACGGAAGTTGGTCGAC 7140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081 TTTACCCAGTTACCCTCCCAAAAAGGAGGTCTCTGAATGGTCAGACGCGAAGTTGGTCAAC 7140

Qy  7141 GGCTACAACCGCTTCCAGCTGCTTCCACTGCCCCCGTACCCTAAGATACGGGGAAGGA 7200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141 GACTACAACCGCTTCCAGCTGCTTCCACTGCCCCCGTACCCTAAGATACGGGGAAGGA 7200

Qy  7201 TTCCACTCAGTCAGCCCCGCCAAACGGCCTACAAAAAGAAGAGTTGGGAAAGAGTGAGTT 7260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201 TTCCACTCAGTCAGCCCCGCCAAACGGCCTACAAAAAGAAGAGTTGGGAAAGAGTGAGTT 7260

Qy  7261 TTCGTGCAGCATGAGCTGACCGACGTGATTAGCTTCTAAAGTGCTTCTAAAGT 7320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261 TTCGTGCAGCATGAGCTGACCGACGTGATTAGCTTCTAAAACTGCTTCTAAAGT 7320

Qy  7321 TCTGTCTGCAACTCGGGACCATCAGTGGTTTCCTCAACAAAGATCATTGGTGTATGT 7380
```

```
Db   7321 TCTGTCTGCAACTCGGCCATCACTAGTGGTTCCTCCAAACAAAGATCATTGGTGTATGT 7380

Qy   7381 GACTGAGCCGCGGGGATCGGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCT 7440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7381 GACTGAGCCGCGGGGATGCGGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCT 7440

Qy   7441 GTTCCCCCATCATACCACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7441 GTTCCCCCATCATACCACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500

Qy   7501 CGGTGTCATGTGGGACTATGATGAAGTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTC 7560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7501 CGGTGTCATGTGGGACTATGATGAAGTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTC 7560

Qy   7561 CCACATCACTGCCCTTCGGGGCACTGATGTTCGTTCTGGAGCAGCCCGCAAGGCGTGTTCT 7620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7561 CCACATCACTGCCCTTCGGGGCACTGATGTTCGTTCTGGAGCAGCCCGCAAGGCGTGTTCT 7620

Qy   7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680

Qy   7681 AGTTCCAAAGGAGGAGAGTCTTCGTGAAGACCCCCAGAGAAACCAACAAAGAAACCCCCAAG 7740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7681 AGTTCCAAAGGAGGAGAGTCTTCGTGAAGACCCCCAGAGAAACCAACAAAGAAACCCCCAAG 7740

Qy   7741 GCTTATCTCGTACCCCCACCTTGAAATGAGATGTGTTGAGAGATGTACTACGGTCAGGT 7800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   7741 GCTTATCTCGTACCCCCACCTTGAAATGAGATGTGTTGAGAGATGTACTACGGTCAGGT 7800

Qy   7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTACGGGTTTGTAGATCCACGTAC 7860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTACGGGTTTGTGACCCACGTAC 7860

Qy   7861 CCGTGTCAAGCGTCTCGTTGTCGATGTGGTCACCCGATGCAGTCGGAGCCACATGCCGATAC 7920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7861 CCGTGTCAAGCGTCTCGTTGTCGATGTGGTCACCCGATGCAGTCGGAGCCACATGCCGATAC 7920

Qy   7921 AGTGTGTTTGACAGTACCATCACACCCGAGAGATATCATGGTGGAGACACATCTACTC 7980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7921 AGTGTGTTTGACAGTACCATCACACCCGAGAGATATCATGGTGGAGACACAGACATCTACTC 7980

Qy   7981 AGCAGCTAAACTCAGTGACCAACACCACCACCGCTGGCATTCACACCATTGCGAGGCAGTTATA 8040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7981 AGCAGCTAAACTCAGTGACCAACACCACCACCGCTGGCATTCACACCATTGCGAGGCAGTTATA 8040

Qy   8041 CGCTGGAGGACCGATGATCGCTTATGATGGCCGAGATCGGATATCGTAGGTGTAGGTC 8100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8041 CGCTGGAGGACCGATGATCGCTTATGATGGCCGAGATCGGATATCGTAGGTGTAGGTC 8100

Qy   8101 TTCCGGGCTCTATACTTGACCTCAAGTTCCAACAGTTTGACCTGCTGCTGAAGGTAAATGC 8160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8101 TTCCGGGCTCTATACTTGACCTCAAGTTCCAACAGTTTGACCTGCTGCTGAAGGTAAATGC 8160

Qy   8161 TGCAGCCGAACAGGCTGGACATGAAGAAGAACCCTCGCCTTCCTTATTTGCGGGCGATGATTGCAC 8220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8161 TGCAGCCGAACAGGCTGGACATGAAGAAGAACCCTCGCCTTCCTTATTTGCGGGCGATGATTGCAC 8220

Qy   8221 CGTAATTGGAAGAGCGCCCGGAGCAGATGCAGAACAAGCAATGCGTGTCTTTGCTAG 8280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8221 CGTAATTGGAAGAGCGCCCGGAGCAGATGCAGAACAAGCAATGCGTGTCTTTGCTAG 8280

Qy   8281 CTGGATGAAGGTGATGGGTGACCACCAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8281 CTGGATGAAGGTGATGGGTGACCACCAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340

Qy   8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400

Qy   8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGTTGCTCTCGCCGAGGGTCTGGG 8460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGTTGCTCTCGCCGAGGGTCTGGG 8460

Qy   8461 ATACAACCCAGTGCTGCGTGGATGGGTATCTAATAATACATCACTACCACCATGTTGTGGGT 8520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8461 ATACAACCCAGTGCTGCGTGGATGGGTATCTAATAATACATCACTACCACCATGTTGTGGGT 8520

Qy   8521 TAGCCGTGTGTTGGCTGTCCATTCATTGGAGACAGATGCTCTTTGAGGACAAACTTCCCGA 8580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8521 TAGCCGTGTGTTGGCTGTCCATTCATTGGAGACAGATGCTCTTTGAGGACAAACTTCCCGA 8580

Qy   8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGGCGCTGTAGAAGATCTGCCCAG 8640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGGCGCTGTAGAGAGATCTGCCCAG 8640

Qy   8641 CATCATTGCTGGTGTGCACGGTGATTGAGGCTTTCTCGGTGGTGTGCGCTACCACCAACGCTGA 8700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8641 CATCATTGCTGGTGTGCACGGTGATTGAGGCTTTCTCGGTGGTGTGCGCTACCACCAACGCTGA 8700

Qy   8701 GATCCTCAGAGTTCCAATCACTAACAGACATGACCATGCCCCCTCGCGAGCCTGGCG 8760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8701 GATCCTCAGAGTTCCAATCACTAACAGACATGACCATGCCCCCTCGCGAGCCTGGCG 8760

Qy   8761 AAAGAAAGCCAGGGCGGTCCTCGCAAGGAGGCGTGCCGGACAGTGGGCGAGCACAGCAAAATT 8820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8761 AAAGAAAGCCAGGGCGGTCCTCGCAAGGAGGCGTGCCGGACAGTGGGCGAGCACACGCAAAATT 8820

Qy   8821 GGCTCGCTTCCTTCTCGGCATGCTCTACATCTAGACCTCTACCAGATTGGATAAGACGAG 8880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8821 GGCTCGCTTCCTTCTCGGCATGCTCTACATCTAGACCTCTACCAGATTGGATAAGACGAG 8880

Qy   8881 CGTGGCTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGGAGGGGATGTGTT 8940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8881 CGTGGCTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGGAGGGGATGTGTT 8940

Qy   8941 TATTACACCACAGAGAAGATTGCAGAAGTTCCTGTGAAGTATTTGGCTGTGTCATTGTTTT 9000
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   8941 TGTTACACCACAGAGAAGATTGCAGAAGTTCCTGTGAAGTATTTGGCTGTGTCATTGTTTT 9000

Qy   9001 TGCCCTAGGGCTCATTGCTGTGTTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA 9060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9001 TGCCCTAGGGCTCATTGCTGTGTTGGACTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA 9060

Qy   9061 CTAACAG----TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAACAGGGGAGACCCC 9116
          ||||||    |||||||||||||||||||||
Db   9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCAGCGCAACAGGGGAGACCCC 9120

Qy   9117 GGGCTTAACGACCCCGC 9133
          |||||||||||||||||
Db   9121 GGGCTTAACGACCCCGC 9137
```

RESULT 12
AR350538

```
LOCUS       AR350538                9143 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 390 from patent US 6586568.
ACCESSION   AR350538
VERSION     AR350538.1  GI:33751681
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 9143)
  AUTHORS   Simons,J.N., Pilot-Matias,T.J., Dawson,G.J., Schlauder,G.G.,
            Desai,S.M., Leary,T.P., Muerhoff,A.S., Erker,J.C., Buijk,S.L. and
            Mushahwar,I.K.
  TITLE     Non-A, non-B, non-C, non-D, non-E hepatitis reagents and methods
            for their use
  JOURNAL   Patent: US 6586568-A 390 01-JUL-2003;
FEATURES             Location/Qualifiers
     source          1..9143
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
```

Query Match   96.4%;   Score 9059.8;   DB 6;   Length 9143;
Best Local Similarity   99.6%;   Pred. No. 0;
Matches 9096;   Conservative 0;   Mismatches 37;   Indels 4;   Gaps 1;

```
Qy      1 ACCACAAACACTCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG   60

Qy     61 CAGGGGCGTGGGGGATTTCCCCTGCCCGGTCTGCAGAGAGGTGGAGCCAACCACCTTAGTAT   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAGGGGCGTGGGGGATTTCCCCTGCCCGGTCTGCAGAGAGGTGGAGCCAACCACCTTAGTAT   120

Qy    121 GTAGGCGCGGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTAGGCGCGGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC   180

Qy    181 CCTGATGGGCGTTCGGTGGTGGTGGTGGCGCTTTAGGCAGCCTTCCAGCCACCA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCTGATGGGCGTTCGGTGGTGGTGGTGGCGCTTTAGGCAGCCCTCCACGCCCCACCA   240

Qy    241 CCTCCCAGATAGAGCGGCGGCACTGTAGGGAGGAAGACCCGGGACCGGTCACTACCAAGGACG   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTCCCAGATAGAGCGGCGGCACTGTAGGGAGGAAGACCCGGGACCGGTCACTACCAAGGACG   300

Qy    301 CAGACCCTCTTTTGAGTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACTTATATGTGT   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGACCCTCTTTTGAGTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACTTATATGTGT   360

Qy    361 TGGGATGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCTTGCGAGGGGAT   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TGGGATGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCTTGCGAGGGGAT   420

Qy    421 CTGGGAGTCTCGTAGACACATGGCCTGTTATTTCTACTCAAACAAGTCCTGTACC   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGGGAGTCTCGTAGACACATGGCCTGTTATTTCTACTCAAACAAGTCCTGTACC   480

Qy    481 TGCGCCCAGAACCGCGCAAGAACAAGCAGCAGGCTTCATATCCTGTGTCCATTAAAAC   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGCGCCCAGAACCGCGCAAGAACAAGCAGCAGGCTTCATATCCTGTGTCCATTAAAAC   540

Qy    541 ATCTGTTGAAAGGGGACAACGAGCAAAGCGCAAAGTCCAGCGCGATGCTCGGCCTCGTAA   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ATCTGTTGAAAGGGGACAACGAGCAAAGCGCAAAGTCCAGCGCGATGCTCGGCCTCGTAA   600

Qy    601 TTACAAAATTGTCGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTACAAAATTGTCGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC   660

Qy    661 TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTCGCAATCTTGGAATCCTTCTCGGA   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTCGCAATCTTGGAATCCTTCTCTGA   720

Qy    721 TTACCCTTTGGGGTGGGTGATTGGTATGTTACAACTCACACAACACCTTCTCTAGTAGGCCCGCTGGT   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TTACCCTTTGGGGTGGGTGATTGGTGATGTTACAACTCACACAACACCTTCTCTAGTAGGCCCGCTGT   780

Qy    781 GGCAGGAGGCGGTCGTTGTTGACCAGTGCGCTGCTGCCAGATACGCTTGCTGGACGATGGAGTCAA   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GGCAGGAGGCGGTCGTTGTTGACCAGTGCGCTGCTGCCAGATGACGCTTGCTGGACGATGGAGTCAA   840

Qy    841 CTTGGGCTACTGGTTCGGTTCGGTGTCCACCTTTTGTGGCTATGTCTGCTCTTTGGGCCTG   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTTGGGCTACTGGTTCGGTTCGGTGTCCACCTTTTGTGGCTATGTCTGCTCTTTGGGCCTG   900

Qy    901 TCCCTGTAGTGGGGGCGCGGGTCACTGACCCAGACACAAATACCACAATTCCTGACCAATTG   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCCCTGTAGTGGGGGCGCGGGTCACTGACCCAGACACAAATACCACAATTCCTGACCAATTG   960

Qy    961 CTGCCAGCGTAATCAGTTATCTATTGTTCTCCTTCCACTTGCCTACACGAGCCTGGGTTG   1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TGGTTCTTGTTGTTACAGGTTCCCAAGGGTTTCACTTGCCTACACGAGCCTGGGTTG   1020

Qy   1021 TGTGATCTGCGCGACGAGTGCTGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA   1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TGTGATCTGCGCGACGAGTGCTGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA   1080

Qy   1081 TTGGACTGGCACGGACTCCTTCTGGCTGACCACCATTGTTGTTATGGGCGCTCTTGT   1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TTGGACTGGCACGGACTCCTTCTGGCTGACCACCATTGTTGTTATGGGCGCGCTTGT   1140

Qy   1141 GACCTGTGACGCCCTTGACATTGGTGAGTTGCGTGTGTCGTGTATTAGTCGGTGACTG   1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTGTGACGCCCTTGACATTGGTGAGTTGCGTGTGTCGTGTATTAGTCGGTGACTG   1200

Qy   1201 GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGA   1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGA   1260

Qy   1261 AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT   1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT   1320

Qy   1321 CGAGGCTGTCATCTTCTTGACCAAACTGCTTCACAAGTACCAAGTACACGCTATTGCCGACTAT   1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CGAGGCTGTCATCTTCTTGACCAAACTGCTTCACAAGTACCAAGTACACGCTATTGCCGACTAT   1380

Qy   1381 GTTTAGCAGTGTACACTACCTGGCGCTCTGACTCATACTACTATGCCTCTCGGGCAA   1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GTTTAGCAGTGTACACTACCTGGCGCTCTGACTCATACTACTATGCCTCTCGGGCAA   1440

Qy   1441 GTGGTATCAGTTGCTCTCTAGAGTATAACAACTCCATATCTTACATAGAAGCGACCCCCAT   1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GTGGTATCAGTTGCTCTCTAGAGTATAACAACTCCATATCTTACATAGAAGCGACCCTCTGG   1500

Qy   1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCCTTTGATGATACCATGTCC   1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCCTTTGATGATACCATGTCC   1560

Qy   1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC   1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC   1620

Qy   1621 CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG   1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG   1680

Qy   1681 TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACACATGGGGTTGCTCGCGTATTCGAA   1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACACATGGGGTTGCTCGCGTATTCGAA   1740

Qy   1741 TGTGCCATCGTACTGCACTATGGGCACTGATGCGAGTGTGGAACGACACTCGCAACACTTA   1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TGTGCCATCGTACTGCACTATGGGCACTGATGCGAGTGTGGAACGACACTCGCAACACTTA   1800

Qy   1801 CGAAGCATGCGGTGTAACACCATGCCAACACCGCACAACGGCTCAGCCCTGAA   1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CGAAGCATGCGGTGTAACACCATGCCAACACCGCACAACGGCTCAGCCCTGAA   1860

Qy   1861 ATTGCTATATTACAATACCCTGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC   1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 ATTGCTATATTACAATACCCTGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC   1920

Qy   1921 AGGCCATTTGTATTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA   1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AGGCCATTTGTATTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA   1980

Qy   1981 TTCCACTCTCCTACCACCGGAGAGGTGGCTAGGTTGCCCGGTACCCCACCTGTGTGTACG   2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TTCCACTCTCCTACCACCGGAGAGGTGGCTAGGTTGCCCGGTACCCCACCTGTGTGTACG   2040

Qy   2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAGG   2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TGGTTCTTGGTTACAGGTTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAGG   2100

Qy   2101 ATTTGATCACCAAAGACAAAGCCTGGAAAAATTATTCAGGTTCTTATATTCCGCACGGTGC   2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   2101 ATTGATCACCAAAGACAAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC 2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TTTGTCTTCTTACGGGAGTTACCACCAAGGCCGTGGTGCTAATTCTGTTGGGGTTGTGTGG 2220

Qy   2161 TTTGTCTTCTTACGGGAGTTACCACCAAGGCCGTGGTGCTAATTCTGTTGGGGTTGTGTGG 2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CAGCAAGTATCTTATTTAGCTGCTACCCTCTGTCCCATCCCAGTCGTATCTCCAAGCTGGA 2280

Qy   2221 CAGCAAGTATCTTATTTAGCTGCTACCCTCTGTCCCATCCCAGTCGTATCTCCAAGCTGGA 2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 TGGTTACCCCTTTGGCTGCGTCCTGTGCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGT 2340

Qy   2281 TGGTTACCCCTTTGGCTGCGTCCTGTGCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGT 2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TTTGTCTAAAGTCAAGTAGCTTCCTTTGCTTGATTTCTTCATCTGTTGCTATCTCCG 2400

Qy   2341 TTTGTCTAAAGTCAAGTAGCTTCCTTTGCTTGATTTCTTCATCTGTTGCTATCTCCG 2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 CTGCAGGCTAGTTATGCTGCCCTTTTAGGGTTTGTGCCATGGCTGCGGGGCTTGCCCCCT 2460

Qy   2401 CTGCAGGCTAGTTATGCTGCCCTTTTAGGGTTTGTGCCATGGCTGCGGGGCTTGCCCCCT 2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 AACTTTCTTTGTTGCAGCAGCTGCTGCCCACCAGATTATGACTGGTGGGTATCTTGCTGCT 2520

Qy   2461 AACTTTCTTTGTTGCAGCAGCTGCTGCCCACCAGATTATGACTGGTGGGTATCTTGCTGCT 2520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 AGTGGCAGGGTTAGTTTTGTGGGCCGCCGCCGTAACCGTGGTCACCGCATAGCTCTGCTTGT 2580

Qy   2521 AGTGGCAGGGTTAGTTTTGTGGGCCGCCGCCGTAACCGTGGTCACCGCATAGCTCTGCTTGT 2580
          |||||||||||||| |||| | |||||||||||||||||||||||||||||||||||||||
Db   2581 AGGTCCTTGGCCTCTGGGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTTCCT 2640

Qy   2581 AGGTCCTTGGCCTCTGGGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTCTCAGC 2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 TTTTGATAACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTAGTCAT 2700

Qy   2641 TTTTGATAACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTAGTCAT 2700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GTCTCGTTTGCTTCACTTGTTTACCTCGTGTGTTTAGTTAACTCCTATCT 2760

Qy   2701 GTCTCGTTTGCTTCACTTGTTTACCTCGTGTGTTTAGTTAACTCCTATCT 2760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 TTGGCAACGTTGGGAGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTCTCCT 2820

Qy   2761 TTGGCAACGTTGGGAGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTCTCCT 2820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 TGTGCTGGTTGTTGTTTCCCCGGTGCACATATGACGGCGTGGTGACTTTCTGTGTGTGTCA 2880

Qy   2821 TGTGCTGGTTGTTGTTTCCCCGGTGCACATATGACGGCGTGGTGACTTTCTGTGTGTGTCA 2880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 CGTAGCTCTTCTATGTTTAAATCCAGTGCAGCATGCGTTCAGTGGGACTGTTCTTAGGGGT 2940

Qy   2881 CGTAGCTCTTCTATGTTTAAATCCAGTGCAGCATGCGTTCAGTGGGACTGTTCTTAGGGGT 2940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 TAGGGCCCATAGAAGTGTCGTCTCGGAAAGTGTCATGCTTCATGGTATTCATTCATTATGT 3000

Qy   2941 TAGGGCCCATAGAAGTGTCGTCTCGGAAAGTGTCATGCTTCATGGTATTCATTCATTATGT 3000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 TCTTAAGTTTTTCCCTCTTAGTGTTTGGTGAGAATGGTGTGTTTTTCTATAAGCACTTGCA 3060

Qy   3001 TCTTAAGTTTTTCCCTCTTAGTGTTTGGTGAGAATGGTGTGTTTTTCTATAAGCACTTGCA 3060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCCATTTTTCCC 3120

Qy   3061 TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCCATTTTTCCC 3120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 TTTTGAAGGGCAAGGCAAGGGGTCTATAGGGAGGAAGACGCTTGGCGTGTGGGGACAC 3180

Qy   3121 TTTTGAAGGGCAAGGCAAGGGGTCTATAGGGAGGAAGACGCTTGGCGTGTGGGGACAC 3180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 GGTTGATGGTTTGCCCGTGTTGCGGCGTCTCGGCGACCTTGTTTTCGCAGGGTTAGCTAT 3240
```

```
Qy   3241 GCCGCCAGATGGGTGGGCCATTACCGCCACCTTTACGCTGCAGTGTCTCTGAACGTGG 3300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 GCCGCCAGATGGGTGGGCCATTACCGCCACCTTTACGCTGCAGTGTCTCTGAACGTGG 3300

Qy   3301 CACGCTGTCAGCGATGGCAGTGCATGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC 3360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 CACGCTGTCAGCGATGGCAGTGCATGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC 3360

Qy   3361 TATCTTCAGATTAGGATCTCTGGCCACTAGCTACATGGATTTGTTTGTGACAACGTGTT 3420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 TATCTTCAGATTAGGATCTCTGGCCACTAGCTACATGGATTTGTTTGTGACACGTGTT 3420

Qy   3421 GTATACTGCTCACCATGGCAGCAAGGGCGCCGGTTGGCTCATCCATCCACAGGCTCTATACA 3480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 GTATACTGCTCACCATGGCAGCAAGGGCGCCGGTTGGCTCATCCATCCCACAGGCTCCATACA 3480

Qy   3481 CCCAATAACCGTTGACGCGGCTAATGCACCAGGACATCTCAACAACCACCACATGTGGAGCTGG 3540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 CCCAATAACCGTTGACGCGGCTAATGCACCAGGACATCTCAACAACCACCACATGTCGAGCTGG 3540

Qy   3541 GTCCCTTACTCGGTGCTCTTGCGGGGGAGACCAAGGGGTATCTGGTAACACGACTGGGGTC 3600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 GTCCCTTACTCGGTGCTCTTGCGGGGGAGACCAAGGGGTATCTGGTAACACGACTGGGGTC 3600

Qy   3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGCGGGGGCCTTCCCAT 3660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGCGGGGGCCTTCCCAT 3660

Qy   3661 GGCTGTTGCCAAGGGTTCTTCAGGTGCCCGATTCTGTGCTCCTCCGGCATGTTATTGG 3720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 GGCTGTTGCCAAGGGTTCTTCAGGTGCCCGATTCTGTGCTCCTCCGGCATGTTATTGG 3720

Qy   3721 GATGTTCACCCGTGCTAGAAATTCTGGCGGTTCAGTCAGTAGTAGCATTAGGGTTAGGCCGTT 3780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 GATGTTCACCCGTGCTAGAAATTCTGGCGGTTCAGTCAGTAGTAGCATTAGGGTTAGGCCGT 3780

Qy   3781 GGTGTGTGCTGGATACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC 3840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 GGTGTGTGCTGGATACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC 3840

Qy   3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATGCCCCCACTGGCAGCGGCAAGTCAAC 3900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATGCCCCCACTGGCAGCGGCAAGTCAAC 3900

Qy   3901 CAAATTACCACTTCTTACATGCAGGAGAAGATGAGGTCTTGGTCCTAAATCCCAGTGT 3960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 CAAATTACCACTTCTTACATGCAGGAGAAGATGAGGTCTTGGTCCTAAATCCCAGTGT 3960

Qy   3961 GGCTACACAGCATCAATGCCAAAGTACAATGCAGCGGCTACGGCGTACGGCGTGAATCAAATTG 4020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 GGCTACACAGCATCAATGCCAAAGTACAATGCAGCGGCTACGGCGTACGGCGTGAATCAAATTG 4020

Qy   4021 CTATTTTAATGCAAATGTACCAACACAGGGGCTTCACTCAGTAATCCTACAGCACATATGGCAT 4080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 CTATTTTAATGCAAATGTACCAACACAGGGGCTTCACTCAGTAATCCTACAGCACATATGGCAT 4080

Qy   4081 GTACCTGACCGGAGCGTCCCGGAACTATGGAATGCCATGCGAATGCCATGC 4140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 GTACCTGACCGGAGCGTCCCGGAACTATGGAATGCCATGCGAATGCCATGC 4140

Qy   4141 TACCGATGCAACCACCGTGTTGGGAGCTCCTGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA 4200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 TACCGATGCAACCACCGTGTTGGGAGCTCCTGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA 4200

Qy   4201 TGTTAGGCTAGTTGTTCTTGCCACGCGCTCACCCCCCCCTGGAGTAATCCCTACACACCACATGC 4260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 TGTTAGGCTAGTTGTTCTTGCCACGCGCTCACCCCCCCCTGGAGTAATCCCTACACACCACATGC 4260

Qy   4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCCACTATCCCCTTTCATGGAAAAAGAT 4320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCCACTATCCCCTTTCATGGAAAAAGAT 4320
```

Right column (positions 5401–6540):

```
Qy  5401  CTTTCTTGGGCCTCATGCGAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGGTTTAGT  5460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401  CTTTCTTGGGCCTCATGCGAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGGTTTAGT  5460

Qy  5461  CACTTTACCTGACAATCCCTTGCATCATGCGTGTTTGCTTTCATTGCGGGTATTACTAC   5520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461  CACTTTACCTGACAATCCCTTGCATCATGCGTGTTTGCTTTCATTGCGGGTATTACTAC   5520

Qy  5521  CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTGGAGGCGCAATTGCGTCCAA   5580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521  CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTGGAGGCGCAATTGCGTCCAA   5580

Qy  5581  GCTTACAGACGCGTAGAGCGCACTGCGTATGACTGGTCATGGTGAGGCGCTGCCGGAACAGCTCT   5640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581  GCTTACAGACGCGTAGAGCGCACTGCGTATGACTGGTCATGGTGAGGCGCTGCCGGAACAGCTCT   5640

Qy  5641  TGGTACATGGACATCGGTTGGGTTTTGTCTTGACATGCCTCTAGGCGGCTAGGCGGCCTCGCCGCCTC   5700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641  TGGTACATGGACATCGGTTGGGTTTTGTCTTGACATGCCTCTAGGCGGCTAGGCGGCCTCGCCGCCTC   5700

Qy  5701  ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGGCCCACTATGGATCAGCT   5760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701  ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGGCCCACTATGGATCAGCT   5760

Qy  5761  TGTCGGTTTAGTCTACTCCGGTTCAATCGGCCGGCCGCAGGAGTTGTGGGCGTCTTGTCGTCAGC   5820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761  TGTCGGTTTAGTCTACTCCGGTTCAATCGGCCGGCCGCAGGAGTTGTGGGCGTCTTGTCGTCAGC   5820

Qy  5821  TTGTGCAATGTTCGCTTTGACAACAGCAGGGCCAGATCACTCGGCCCAACAGACTTCTTAC   5880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821  TTGTGCAATGTTCGCTTTGACAACAGCAGGGCCAGATCACTCGGCCCAACAGACTTCTTAC   5880

Qy  5881  TATGCTTGCTAGGAGCAACACTGTATGTAATAGATACTTTATTGCCACTCGTGACATCCG   5940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881  TATGCTTGCTAGGAGCAACACTGTATGTAATAGATACTTTATTGCCACTCGTGACATCCG   5940

Qy  5941  CAGGAAGATACTGGCATTCTGACGCAGCATCTACCCCCTGGAGTGTCATATCAGCTGCAT   6000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941  CAGGAAGATACTGGCATTCTGACGCAGCATCTACCCCCTGGAGTGTCATATCAGCTGCAT   6000

Qy  6001  CCGTTGGGCTCCACACCCCGACGGAGGATGAGTTGCGGCCTCATTGTCCCAAGCACGCGTTCATTGCTTGAGAT   6060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001  CCGTTGGGCTCCACACCCCGACGGAGGATGAGTTGCGGCCTCATTGTCCCAAGCACGCGTTCATTGCTTGAGAT   6060

Qy  6061  TTGGCAGTATGTGCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGCTGGAGTTCA   6120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061  TTGGCAGTATGTGCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGCTGGAGTTCA   6120

Qy  6121  GAGCATGGTTAACATTCCTGGTTGTCCTTCTACAGCTGCCAGAGAGGGTACAAGGGCCC   6180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121  GAGCATGGTTAACATTCCTGGTTGTCCTTCTACAGCTGCCAGAGAGGGTACAAGGGCCC   6180

Qy  6181  CTGGATTGGATCAGTATGCGGTTAGGGACTACTGTAAATATGGGAACTCATCTTTTC   6240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181  CTGGATTGGATCAGTATGCGGTTAGGGACTACTGTAAATATGGGAACTCATCTTTTC   6240

Qy  6241  TGTTGAGAATGGTTTGCAAAACTTACAAAGGACCCCAGAACTTGTTCAAATTACTGGAG   6300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241  TGTTGAGAATGGTTTGCAAAACTTACAAAGGACCCCAGAACTTGTTCAAATTACTGGAG   6300

Qy  6301  AGGGGCTGTTCCAGTCAACGCGTAGGCTGTGTGGCGGCTAGACCCGGACCCCAACCGATTG   6360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301  AGGGGCTGTTCCAGTCAACGCGTAGGCTGTGTGGCGGCTAGACCCGGACCCCAACCGATTG   6360

Qy  6361  GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGGGAAATGGGAGA   6420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361  GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGGGAAATGGGAGA   6420

Qy  6421  TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGTGCCCCCAAC   6480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421  TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGTGCCCCCAAC   6480

Qy  6481  CTTGAGAGCTGCAGCGCCGTGACGGCGTCGTATCGTAGGTTCAGTGTTACTAGGTGAGCCTTAGT
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Left column (positions 4321–5460):

```
Qy  4321  TAAGGAGGAAAAATCTGAAGAAAGGGAGAAACACCTTATCTTTGAGGCTACCAAAAAACACTG   4380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321  TAAGGAGGAAAAATCTGAAGAAAGGGAGAAACACCTTATCTTTGAGGCTACCAAAAAACACTG   4380

Qy  4381  TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGAATAACAGCTGTCTCTTACTATAGGGG   4440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381  TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGAATAACAGCTGTCTCTTACTATAGGGG   4440

Qy  4441  ATGTGACATCTCAAAAATCCCTGAGGGCGACTGTGTAGTGTCCACTGAGTGCCACTGCCTTGTG   4500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441  ATGTGACATCTCAAAAATCCCTGAGGGCGACTGTGTAGTGTCCACTGAGTGCCACTGCCTTGTG   4500

Qy  4501  TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTGCAGCCTCATGTGGTAGAAGGCAC   4560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501  TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTGCAGCCTCATGTGGTAGAAGGCAC   4560

Qy  4561  ATGCCATTGTTGACCCTACTTTCACCATGTGTTCGTGTGTGGGTTTCAGC   4620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561  ATGCCATTGTTGACCCTACTTTCACCATGTGTTCGTGTGTGGGTTTCAGC   4620

Qy  4621  AATAGTTAAAGGCCAGCGTAGGGGCCGCACAGCGCCGTTGGGAGAGCTGCATATACTACTA   4680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621  AATAGTTAAAGGCCAGCGTAGGGGCCGCACAGCGCCGTTGGGAGAGCTGCATATACTACTA   4680

Qy  4681  TGTAGACGGGAGTTGTACCCCCTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT   4740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681  TGTAGACGGGAGTTGTACCCCCTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT   4740

Qy  4741  CGACGCAGCCAAGGCATGGTATGGTTTGTCATCAACAGAAGTCAAACTATTCTGACAC   4800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741  CGACGCAGCCAAGGCATGGTATGGTTTGTCATCAACAGAAGTCAAACTATTCTGACAC   4800

Qy  4801  CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGTCAAATTTGGACGAGTGGCTGATCT   4860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801  CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGTCAAATTTGGACGAGTGGCTGATCT   4860

Qy  4861  CTTTTCTATGGTCAACCCCGAACCTTCATTTGTCAATACTGCAAAAAGAAGAACTGCTGCAA   4920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861  CTTTTCTATGGTCAACCCCGAACCTTCATTTGTCAATACTGCAAAAAGAAGAACTGCTGCAA   4920

Qy  4921  TTATGTTTGTTGTTGACTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTACTGCTCC   4980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921  TTATGTTTGTTGTTGACTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTACTGCTCC   4980

Qy  4981  CAATGAGCCACCACGGTGGCAGGAGCCCGGCTGGGAAAAAACCTTGTGTGGGGTTTCTGTG   5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981  CAATGAGCCACCACGGTGGCAGGAGCCCGGCTGGGAAAAAACCTTGTGTGGGGTTTCTGTG   5040

Qy  5041  GCGCTTGACGCGCGACTGTCCCCACGGTGCTACTGGTTGACGAAGAGAAATCGT   5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041  GCGCTTGACGCGCGACTGTCCCCACGGTGCTACTGGTTGACGAAGAGAAATCGT   5100

Qy  5101  AATGTGCTTCACTGAAGTCATCAATATCCCTTGGAGGCCATGGCTGTTGCTGCAAGGTGAA   5160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101  AATGTGCTTCACTGAAGTCATCAATATCCCTTGGAGGCCATGGCTGTTGCTGCAAGGTGAA   5160

Qy  5161  GGCTATGCTTATCTAGCCATTGACACTTTGAGGCCCCATTGGAAAACCGCCCTTGAAAACTTAACAC   5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161  GGCTATGCTTATCTAGCCATTGACACTTTGAGGCCCCATTGGAAAACCGCCCTTGAAAACTTAACAC   5220

Qy  5221  TATTACATCAGTCCGGTGCTACTGGCTACTGGTTGACGAAGAGAAATCGT   5280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221  TATTACATCAGTCCGGTGCTACTGGCTACTGGTTGACGAAGAGAAATCGT   5280

Qy  5281  GGAGGAGTTGTGCATCATTCATTCCCTTGGAGGCCATGGCCATGGTTGCAAGCTGCTGTGCAATTGACAAGCTGAA   5340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281  GGAGGAGTTGTGCATCATTCATTCCCTTGGAGGCCATGGCCATGGTTGCAAGCTGCTGTGCAATCGATAAGCTGAA   5340

Qy  5341  GAGTACAATCAACCACCACAGTTCCTTTCACATTGAAAACCGCCCTTGAAAATGCTATCATAGAGTATTGCTGTGGTTTAGT   5400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341  GAGTACAATAAACCACCACAGTTCCTTTCACATTGAAAACCGCCCTTGAAAACTTAACAC   5400
```

us-10-009-002-1.rge

```
Db  6481 CTTGAGAGCTGCAGTGGCCGTGGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540

Qy  6541 AACTCCTTGGACGACATCTGCTTGCTGTTACGGTCCTGACGCGTAAGCGTAAAACTGTTAA 6600
         ||||||||||||||||||||||||||||||||||||  |||||||||||||||||||||||
Db  6541 AACTCCTTGGACGACATCTGCTTGCTGTTACGGTCCCGACGCGTAAGCGTAAAACTGTTAA 6600

Qy  6601 GCTTCCCTTCCGCGTTGACGGTTCACACCGGTTGCGCACACAACAACTCCACAAACACTGGT 6660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601 GCTTCCCTTCCGCGTTGACGGTTCACACCGGTTGCGCACACAACAACTCCATAAACACTGGT 6660

Qy  6661 TGCACTTGAGACAACAAATGACTGTAATTCCACAAACAACACTCCTAGTGATGAAGCCGCAGT 6720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661 TGCACTTGAGACAACAAATGACTGTAATTCCACAAACAACACTCCTAGTGATGAAGCCGCAGT 6720

Qy  6721 GTCCGGCTCTTGTTTCAAACAGGAGTTGCGGGTACAAACCAATTGCTTGAGGCAATTTC 6780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721 GTCCGGCTCTTGTTTTCAAACAGGAGTTGCGGGTACAAACCAATTGCTTGAGGCAATTTC 6780

Qy  6781 AGGAGTGTCATGTCCTGAAAGCTGCACCGCCAAACTGCCAGCCCCTCCATCGAAGAGGTAGTGGTAAGAAA 6840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781 AGGAGTGTCATGTCCTGAAAGCTGCACCGCCAAACTGCCAGCCCCTCCATCGAAGAGGTAGTGGTAAGAAA 6840

Qy  6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCCTCCCCCCTCGCCTGCCCGAGATCCGTCCC 6900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCCTCCCCCCTCGCCTGCCCGAGATCCGTCCC 6900

Qy  6901 AGGAGTGTCATGTTCCTGAAAGCGCTGAAGCCTGAACGAAGTGACCCGGTTAGAGAGGTCCTTCAAACCT 6960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901 AGGAGTGTCATGTTCCTGAAAGCGCTGAAGCCTGAACGAAGTGACCCGGTTAGAGAGGTCCTTCAAACCT 6960

Qy  6961 CCCTCCTTCACCACTGTTCTACAGTTGGCTGGCATGCCGGATGGCGGGTTGGGCCCTGTTGGGAGGCGGGTGA 7020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961 CCCTTCTTCACCACTGTTCTACAGTTGGCTGGCATGCCGGATGGCGGGTTGGGCCCTGTTGGGAGGCAGTGA 7020

Qy  7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGAAGACCGCCGAAATGACCGAAAACAGGCGGGAAAGGA 7080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGAAGACCGCCGAAATGACCGAAAACAGGCGGGAAAGGA 7080

Qy  7081 TTTACCCAGTTACCCTCCAGCGCTTCAGCCCCCCCCCAAACGGCCTACTCGGTAGTAGACGGGGAAAGGA 7140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081 TTTACCCAGTTACCCTCCAGCGCTTCAGCCCCCCCCCAAACGGCCTACTCGGTAGTAGACGGGGAAAGGA 7140

Qy  7141 GGCTACAACCGCTTCAGCCCGTTCAGCTGGACCGACGACTGGATTAGCTTCAAACAAAGATCATTGGTGTATGT 7200
         |||||||||||||||||||||||||||||||||||||||                        
Db  7141 GACTACAACCGCTTCAGCCCGTTCAGCTGGACCGACGACTGGATTAGCTTCAAACAAAGATCATTGGTGTATGT 7200

Qy  7201 TTCCACTCAGTCAGCCGGGATGCGGGGATCGCGGGGATCCGCCGCCAAACGGCCTTAGAGAAGAAAAGTTGGGGAAAGGA 7260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201 TTCCACTCAGTCAGCCGGGATGCGGGGATCGCGGGGATCCGCCGCCAAACGGCCTTAGAGAAGAAAAGTTGGGGAAAGGA 7260

Qy  7261 TTCGTGCAGCATGATGAGCTCATCGGGGCCATCGGGGCCGACGACGTGATTAGCTTCAAACAAAGATCGCTTCTAAAGT 7320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261 TTCGTGCAGCATGATGAGCTCATCGGGGCCATCGGGGCCGACGACGTGATTAGCTTCAAACAAAGATCGCTTCTAAAGT 7320

Qy  7321 TCTGTTCTGCAACTCGGGGCCATCGGGGCCGCCCAGTACTTGGTTCCTCAAACAAAGATCATTGGTGTATGT 7380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7321 TCTGTTCTGCGACTCGGGGCCATCGGGGCCGCCCAGTACTTGGTTCCTCAAACAAAGATCATTGGTGTATGT 7380

Qy  7381 GACTGAGCGCGCGGGATGCGGGGATCGCGGGGATTGGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCT 7440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381 GACTGAGCGCGCGGGATGCGGGGATCGCGGGGATTGGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCT 7440

Qy  7441 GTTCCCCCATCATACCACCACCACAGCGAGATTGGCTAAGGAAAAAGCTTCAAAGTGT 7500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441 GTTCCCCCATCATACCACCACCACAGCGAGATTGGCTAAGGAAAAAGCTTCAAAGTGT 7500

Qy  7501 CGGTGTCATGTGGGACTATGTGGGACTAGCAGCTAGCAGCGACGCTCACACGCGCCACCCACCCCTCAAGTC 7560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7501 CGGTGTCATGTGGGACTATGTGGGACTAGCAGCTAGCAGCGACGCTCACACGCGCCACCCACCCCTCAAGTC 7560

Qy  7561 CCACATCACTGGCCTTCGGGGCACTGCTCGGTTCGTTCTCTGGAGCACGCCCAAGGCTGTTCT 7620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7561 CCACATCACTGGCCTTCGGGGCACTGCTCGGTTCGTTCTCTGGAGCACGCCCAAGGCTGTTCT 7620
```

```
Db  7621 GGACTTGCAGAGAAGTGTGTCGAGGCAGAGTGTCGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680

Qy  7621 GGACTTGCAGAGAAGTGTGTCGAGGCAGAGTGTCGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7621 GGACTTGCAGAGAAGTGTGTCGAGGCAGAGTGTCGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680

Qy  7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAACCAACAAAGAAACCCCCAAG 7740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAACCAACAAAGAAACCCCCAAG 7740

Qy  7741 GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGT 7800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7741 GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGT 7800

Qy  7801 TGCTCCTGACGTAGTAGTAAAGCTGTCATGGAGTGCTCACGGGTTGTAGATCCACGTAC 7860
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7801 TGCTCCTGACGTAGTAGTAAAGCTGTCATGGAGTGCTCACGGGTTTGTCGACCACGTAC 7860

Qy  7861 CCGTGTCAAGCGTCTGTGTGTCAGTGCATGCCCGATGCGCAGTCGATCGGAGGCACATGCGATAC 7920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7861 CCGTGTCAAGCGTCTGTGTGTCAGTGCATGCCCGATGCGCAGTCGATCGGAGGCACATGCGATAC 7920

Qy  7921 AGTGTGTTTGACAGTACCATCACACCCGAGATATCATGGTGGAGACAGACATCTACTC 7980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7921 AGTGTGTTTGACAGTACCATCACACCCGAGATATCATGGTGGAGACAGACATCTACTC 7980

Qy  7981 AGCAGCTAAACTCAGTGACCAACACCGATGCGCATTCACACCAGCTGCGAGGCAGTTATA 8040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7981 AGCAGCTAAACTCAGTGACCAACACCGATGCGCATTCACACCAGCTGCGAGGCAGTTATA 8040

Qy  8041 CGCTGGGAGACCGATGATCGCTTATGGCGGATCGGAGATCGGAGATCGTAGGTGTAGGTC 8100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8041 CGCTGGGAGACCGATGATCGCTTATGGCGGATCGGAGATCGGAGATCGTAGGTGTAGGTC 8100

Qy  8101 TTCCGGCGTCTATACTACCCAAGTTCAACAGTTGACCTGCTGGCTGAAGGTAAATGC 8160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8101 TTCCGGCGTCTATACTACCCAAGTTCAACAGTTGACCTGCTGGCTGAAGGTAAATGC 8160

Qy  8161 TGCAGCCGAACAGGCTGGCATGAGAACCCTCGCCTTATTGCGGCGATGATTGCAC 8220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161 TGCAGCCGAACAGGCTGGCATGAGAACCCTCGCCTTATTGCGGCGATGATTGCAC 8220

Qy  8221 CGTAATTGGAAGAGCGCCGGAACAGCAGATGCAGACAAACAAGCAATGCGTCTTTGCTAG 8280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8221 CGTAATTGGAAGAGCGCCGGAACAGCAGATGCAGACAAACAAGCAATGCGTCTTTGCTGA 8280

Qy  8281 CTGGATGAAGGTGATGGGTGCGACCCACAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281 CTGGATGAAGGTGATGGGTGCGACCCACAGATTGTGTGCCTCAACCCAAATACAGTTTTGA 8340

Qy  8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400

Qy  8401 CTACTTCTTACAAGAGATCCTCGTATCCCCCTGGCAGGTGCTGTCTCGGAGGGTCTCGGG 8460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8401 CTACTTCTTACAAGAGATCCTCGTATCCCCCTGGCAGGTGCTGTCTCGGAGGGTCTCGGG 8460

Qy  8461 ATACAACCCCAGTGCTGCCGTGATTGGTGTATCCCATCCCATGGGTGTTGTGGGGT 8520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8461 ATACAACCCCAGTGCTGCCGTGATTGGTGTATCCCATCCCATGGGTGTTGTGGGGT 8520

Qy  8521 TAGCCGTGTGTGGCTCGTCCATTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8521 TAGCCGTGTGTGGCTCGTCCATTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580

Qy  8581 GACTGTGTCCGGTATGGTATGGGAAAATTATACGGTGCCTGTAGAAGATCGCGCCCAG 8640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8581 GACTGTGTCCGGTATGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCGCGCCCAG 8640

Qy  8641 CATCATTGCGTGTGCACGGTTCGTTCTCCGGTGTGCGCTACCAACGCTGA 8700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8641 CATCATTGCGTGTGCACGGTTCGTTCTCCGGTGTGCGCTACCAACGCTGA 8700
```

```
Qy        CCTGATGGGCGTTCATGGGTTCGGTGGTGGTGGTGGGAAGACCGGGACCGGCCTTTAGGCAGCCTCACGCCACCA  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8701  GATCCTCAGAGTTCCAATCACTAACAGACATGACCATGCCCCCCTGCCGCGGCG                         8760

Qy  241   CCTCCCAGATAGAGCGGCGGCACTGTAGGGAAGACCGGGACCGGGTCACTACCAAGGACG  300
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  8701  GATCCTCAGAGTTCCAATCACTAACAGACATGACCATGCCCCCCTGCCGCGGCG  8760

Qy  241   CCTCCCAGATAGAGCGGCGGCACTGTAGGGAAGACCGGGACCGGGTCACTACCAAGGACG  300
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  8761  AAAGAAAGCCAGGGCGTCCTGCCAGCGGCCGTGGCGGAGCCGGCACACGCAAATT  8820

Qy  301   CAGACCTCTTTTTGAGTATCACGGCCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTGT  360
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  8761  AAAGAAAGCCAGGGCGTCCTGCCAGCGGCCGTGGCGGAGCCGGCACACGCAAATT  8820

Qy  301   CAGACCTCTTTTTGAGTATCACGGCCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTGT  360
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  8821  GGCTCGGCGTTCCTTCTCTCGGCATGCTACATCTAGACCTCTACCAGATTTGGATAAGACAGAG  8880

Qy  361   TGGGATGGTTGGGGTTAGCCTTCATCATACCGTCGCCTCGATAGGGTCCTTGCGAGGGGAT  420
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  8821  GGCTCGGCGTTCCTTCTCTCGGCATGCTACATCTAGACCTCTACCAGATTTGGATAAGACAGAG  8880

Qy  361   TGGGATGGTTGGGGTTAGCCTTCATCATACCGTCGCCTCGATAGGGTCCTTGCGAGGGGAT  420
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  8881  CGTGGCTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGGAGGGGGGATGTGTT  8940

Qy  421   CTGGGAGTCTCGTAGACCGTAGCACATGCCGTGTATTCTACTCAAACAAGTCCTGTACC  480
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  8881  CGTGGCTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGGAGGGGGGATGTGTT  8940

Qy  421   CTGGGAGTCTCGTAGACCGTAGCACATGCCGTGTATTCTACTCAAACAAGTCCTGTACC  480
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  8941  TATTACACCACAGAGAGAAGAGAGATTGCAGAGAAGAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTT  9000

Qy  481   TGCCCCAGAACCGGCAAGAACAGACAGAGCGCAAAGGCAAAGTCAGCGGCGATGCTCGGCGCTCGGCCTCGTAA  540
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  8941  TGTTACACCACAGAGAGAAGAGAGATTGCAGAGAAGAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTT  9000

Qy  481   TGCCCCAGAACCGGCAAGAACAGACAGAGCGCAAAGGCAAAGTCAGCGGCGATGCTCGGCGCTCGGCCTCGTAA  540
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  9001  TGCCCTAGGGCGTTCATTGCTGTGTTGGATTAGCCTCAGCTGAACCCCCCAAATTCAAAATTAA  9060

Qy  541   ATCTGTTGAAAGGGGACAACGAGCAAAGGCAAAGTCAGCGGCGATGCTCGGCGCTCGGCCTCGTAA  600
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  9001  TGCCCTAGGGCGTTCATTGCTGTGTTGGATTAGCCTCAGCTGAACCCCCCAAATTCAAAATTAA  9060

Qy  541   ATCTGTTGAAAGGGGACAACGAGCAAAGGCAAAGTCAGCGGCGATGCTCGGCGCTCGGCCTCGTAA  600
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  9061  CTAACAG----TTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGCAACAGGGGAGACCCC  9116

Qy  601   TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGCAATCTCGCCAATCTTGGAATCCTTCTGGA  720
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  9061  TTAACAGTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC  9120

Qy  601   TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGCAATCTCGCCAATCTTGGAATCCTTCTGGA  720
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  9117  GGGCTTAACGACCCCGC  9133

Qy  661   TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTCGCCAATCTTGGAATCCTTCTGGA  720
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  9121  GGGCTTAACGACCCCCGC  9137

Qy  661   TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTCGCCAATCTTGGAATCCTTCTGGA  720

Qy  721   TTACCCTTTGGGGTGGATTGGTGATGTTACAACTCACACACCCTCAGTAGGCCCGCTGTGT  780
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  721   TTACCCTTTGGGGTGGATTGGTGATGTTACAACTCACACACCCTCAGTAGGCCCGCTGGT  780

Qy  781   GGCAGGAGCGGGTCGTTCGACCAGTCTGCGCCAGTAGTAGCGTCTTGCTGGAGGATGGAGTCAA  840
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  781   GGCAGGAGCGGGTCGTTCGACCAGTCTGCGCCAGTAGTAGCGTCTTGCTGGAGGATGGAGTCAA  840

Qy  841   CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTGTGGTGTCGTATGTGCTGCTATCTTTGGCCTG  900
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  841   CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTGTGGTGTCGTATGTGCTGCTATCTTTGGCCTG  900

Qy  901   TCCCTGTAGTGGGGCGGCGGGTCACTGACCGGCACACAAATACCACAAATACCGACCAATTG  960
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  901   TCCCTGTAGTGGGGCGGCGGGTCACTGACCGGCACACAAATACCACAAATACCGACCAATTG  960

Qy  961   CTGCCAGCGTAATCAGGTTATCATTGTCTCCTTCCACTTGCCACTGCCTACAGAGCCTGGTTG  1020
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  961   CTGCCAGCGTAATCAGGTTATCATTGTCTCCTTCCACTTGCCACTGCCTACAGAGCCTGGTTG  1020

Qy  1021  TGTGATCTGTGCGACGAGCGTGCTGGCTTCCCGCCAATCCGTACATCTCACACCGTTCCAA  1080
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  TGTGATCTGCGCGGACGAGTGCTGCTGGCTTCCCGCCAATCCGTACATCTCACACCCTTCCAA  1080

Qy  1081  TTGGACTGGCACGGACTCCTTCGTCGGCTAGGAATGCTCCTGGAGCACCCCCCTAG  1140
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  TTGGACTGGCACGGACTCCTTCGTCGGCTAGGAATGCTCCTGGAGCACCCCCCTAG  1140

Qy  1141  GACCTGTGACGCCCTTGACATTGGTGAGTTGTGTGGTGTATTAGTTCGGTGACTG  1200
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  GACCTGTGACGCCCTTGACATTGGTGAGTTGTGTGGTGTATTAGTTCGGTGACTG  1200

Qy  1201  GCTTGTCAGGCACTGCTTATTCACATAGACCTGAAACTGGTACTTGTTACCTGGA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GCTTGTCAGGCACTGCTTATTCACATAGACCTGAAACTGGTACTTGTTACCTGGA  1260

Qy  1261  AGTGCCACTGGAATAGATCCTGGGTTCCGTGGTGGTGTGGCGGTTATCGGGTGATGCGCCGGCAAGGT  1320
          ||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 13

AR350541

LOCUS	AR350541	9143 bp	DNA	linear	PAT 17-AUG-2003

DEFINITION Sequence 393 from patent US 6586568.
ACCESSION   AR350541
VERSION   AR350541.1   GI:33751684
KEYWORDS   .
SOURCE   Unknown.
ORGANISM   Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 9143)
AUTHORS   Simons,J.N., Pilot-Matias,T.J., Dawson,G.J., Schlauder,G.G.,
            Desai,S.M., Leary,T.P., Muerhoff,A.S., Erker,J.C., Buijk,S.L. and
            Mushahwar,I.K.
TITLE   Non-A, non-B, non-C, non-D, non-E hepatitis reagents and methods
            for their use
JOURNAL   Patent: US 6586568-A 393 01-JUL-2003;
FEATURES   Location/Qualifiers
source   1..9143
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN

Query Match   96.4%; Score 9059.8; DB 6; Length 9143;
Best Local Similarity   99.6%; Pred. No. 0;
Matches 9096; Conservative   0; Mismatches   37; Indels   4; Gaps   1;

```
Qy  1    ACCACAAACACTCCAGTTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCCTAG  60
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1    ACCACAAACACTCCAGTTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCCTAG  60

Qy  61   CAGGGCGTGGGGGATTCCCCTGCCCGTCTGCAGAAGGGTGAGCCAACCAACCACCTTAGTAT  120
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  61   CAGGGCGTGGGGGATTCCCCTGCCCGTCTGCAGAAGGGTGAGCCAACCAACCACCTTAGTAT  120

Qy  121  GTAGGCGGCGGGGACTCATGACGCGTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GTAGGCGGCGGGGACTCATGACGCGTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180

Qy  181  CCTGATGGGCGTTCATGGGTTCGGTGGTGGTGGGCGGTTTAGGCAGCCTCACGCCACCA  240
```

Db

Qy   1261 AGTGCCCACTCGAATAGATCCTGGGTTCGTAGGGTTTATCGGGTGGATGGCCGGCAAGGT 1320
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db

Qy   1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCGCTATTGCGACTAT 1380
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCGCTATTGCGACTAT 1380

Qy   1381 GTTTAGCAGTGTACACTACCTGGCGCGGTTGGCGCTTCTGATCTACTATAGCCTCTCGGGGCAA 1440
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GTTTAGCAGTGTACACTACCTGGCGCGGTTGGCGCTTCTGATCTACTATAGCCTCTCGGGGCAA 1440

Qy   1441 GTGGTATCAGTTGCTTCCTAGCGGCTTATGCTTACATAGAAGCGACCTTGTACAGTCCAAAGTGGAC 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GTGGTATCAGTTGCTTCCTAGCGGCTTATGCTTACATAGAAGCGACCTTGTACAGTCCAAAGTGGAC 1500

Qy   1501 CAGGGTGCCCACTGGAGTCTCAATAGCTGAGTTTGCTCGCCTTTGATGATACCATGTCC 1560
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CAGGGTGCCCACTGGAGTCTCAATAGCTGAGTTTGCTCGCCTTTGATGATACCATGTCC 1560

Qy   1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTTGTTACAGTCCAAAGTGGAC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTTGTTACAGTCCAAAGTGGAC 1620

Qy   1621 CAGGCCTATCACTCTAGAGTAGTAACAACTCCATATCTTGGTACCCCTATACAAATCCTGG 1680
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CAGGCCTGTCACTCTAGAGTAGTAACAACTCCATATCTTGGTACCCCTATACAAATCCTGG 1680

Qy   1681 TGCGAGGGATGTATGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA 1740
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGCGAGGGATGTATGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA 1740

Qy   1741 TGTGCCATCGTACTGCACTATGGGCACTGATGCAGTGTGAACGACACTCGCAACACTTA 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TGTGCCATCGTACTGCACTATGGGCACTGATGCAGTGTGAACGACACTCGCAACACTTA 1800

Qy   1801 CGAAGCATGCCGTGTAACAACATGGCTAACACCGGCTCAGGCCTCAGCCCTGAA 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CGAAGCATGCCGTGTAACAACATGGCTAACACCGGCTCAGGCCTCAGCCCTGAA 1860

Qy   1861 ATTGGCTATATTACAATACCCTGGGTCTGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 ATTGGCTATATTACAATACCCTGGGTCTGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC 1920

Qy   1921 AGGCCATTTGTATTTGAGGGATCAGATACCCCTATAGTTTACTTTATGACCCTGTGAA 1980
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AGGCCATTTGTATTTGAGGGATCAGATACCCCTATAGTTTACTTTATGACCCTGTGAA 1980

Qy   1981 TTTCCACTCTCCTCCACCGGGAGAGGTGGGCTAGGTGCCCGGTACCCACCTGTGTACG 2040
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TTTCCACTCTCCTCCACCGGGAGAGGTGGGCTAGGTGCCCGGTACCCACCTGTGTACG 2040

Qy   2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTTACGTGATGTGAAAGACCTAGCCACAGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTTACGTGATGTGAAAGACCTAGCCACAGG 2100

Qy   2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGGCCACGGGTGC 2160
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGGCCACGGGTGC 2160

Qy   2161 TTTGTCTTACGGGAGTTACCACCAAGGGCCGTGGCTCTAATTCTGTTGGGGTTGTGTGG 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TTTGTCTTACGGGAGTTACCACCAAGGGCCGTGGCTCTAATTCTGTTGGGGTTGTGTGG 2220

Qy   2221 CAGCAAGTAGTATCTTATTTTAGCCTACCTCTGTACTTGTCCCCTTTGTTTTGGGGCGCGTTC 2280
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CAGCAAGTAGTATCTTATTTTAGCCTACCTCTGTACTTGTCCCCTTTGTTTTGGGGCGCGTTC 2280

Qy   2281 TGGTTACCCCTTTGCGTCCTGTCTCCATCCCAGTCGTATCCCAGTCGTATCCCAAGCTGGCTGGGATGT 2340
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 TGGTTACCCCTTTGCGTCCTGTCTCCATCCCAGTCGTATCCCAGTCGTATCCCAAGCTGGCTGGGATGT 2340

Qy   2341 TTTGTCTAAAGCTCAAGTAGCTCCCTTTGCTTTGATTTTCTTCATCTGTTGCTATCCCG 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TTTGTCTAAAGCTCAAGTAGCTCCCTTTGCTTTGATTTTCTTCATCTGTTGCTATCCCG 2400

Qy   2401 CTGCAGGCTACGTTATGCTGCCCATGCGGCCATGCGGCCGCGCGGGGCTTGCCCCT 2460
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 CTGCAGGCTACGTTATGCTGCCCATGCGGCCATGCGGCCGCGCGGGGCTTGCCCCT 2460

Qy   2461 AACTTTCTTGTTGTGCAGCAGCTGCGCTGCCAACCAGATTATGACTGGTGGTGCGGACTGCT 2520
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 AACTTTCTTGTTGTGCAGCAGCTGCGCTGCCAACCAGATTATGACTGGTGGTGCGGACTGCT 2520

Qy   2521 AGTGGCAGGGTTAGTTTTGTGTGGCGCCGTAACCGTGGTCACCGCATAGCTCTGCTTGT 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 AGTGGCAGGGTTAGTTTTGTGTGGCGCCGTAACCGTGGTCACCGCATAGCTCTGCTTGT 2580

Qy   2581 AGGTCCTTGGCCTCTGGTAGGCCTTTTAACCCTCCTTGCATTGGTTACGCCTGCTTCAGC 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 AGGTCCTTGGCCTCTGGTAGGCCTTTTAACCCTCCTTGCATTGGTTACGCCTGCTTCAGC 2640

Qy   2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 TTTTGACACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT 2700

Qy   2701 GTCTCGTTTGGCTTCTTGCTCACTTGTTACCTCGCTGTGCTCTAAGACCGGAGAGTTTTCCT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GTCTCGTTTGGCTTCTTGCTCACTTGTTACCTCGCTGTGCTCTAAGACCGGAGAGTTTTCCT 2760

Qy   2761 TTGGCAACGTTGGGAGAATTGGTTTCCCGGTGCGACATATGACGCGCTGGTGACTTTCTGTGTGTCA 2820
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 TTGGCAACGTTGGGAGAATTGGTTTCCCGGTGCGACATATGACGCGCTGGTGACTTTCTGTGTGTCA 2820

Qy   2821 TGTGTCTGGTTGTTTCCCCGGTGCGTGCGTGGTGCTGGTGACCGCTGGTGACTTTCTGTGTGTCA 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 TGTGTCTGGTTGTTTCCCCGGTGCGTGCGTGGTGCTGGTGACCGCTGGTGACTTTCTGTGTGTCA 2880

Qy   2881 CGTAGCTCTTCATGTTTAACATCCAGTGCAGCATCGTTCTTGGGACTGACTCTAGGGT 2940
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 CGTAGCTCTTCATGTTTAACATCCAGTGCAGCATCGTTCTTGGGACTGACTCTAGGGT 2940

Qy   2941 TAGGGCCCATAGAATGTTGGTGCGTCTCGGAAAGTGTCATGCGTATTCTCATTATGT 3000
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 TAGGGCCCATAGAATGTTGGTGCGTCTCGGAAAGTGTCATGCGTATTCTCATTATGT 3000

Qy   3001 TCTTAAGTTTTTCCTCTTAGTGTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA 3060
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 TCTTAAGTTTTTCCTCTTAGTGTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA 3060

Qy   3061 TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCAAGAGCCATTTTTCCC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCAAGAGCCATTTTTCCC 3120

Qy   3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGGTCTATAGGAGGAAGCGCTTGGGCGTGTGGGGACAC 3180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGGTCTATAGGAGGAAGCGCTTGGGCGTGTGGGGACAC 3180

Qy   3181 GGTTGATGGTTGTTGCCCGGTGTGTTGCCCGGTGTGTTTTCGGGGGTTGGCTAT 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 GGTTGATGGTTGTTGCCCGGTGTGTTGCCCGGTGTGTTTTCGGGGGTTAGCTAT 3240

Qy   3241 GCCGCCAGATGGGTGGCCATTACCGACCTTTTACGCGTCGCAGTGTCTCTGAACGTGG 3300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 GCCGCCAGATGGGTGGGCCATTACCGACCACCTTTTACGCGTCGCAGTGTCTCTGAACGTGG 3300

Qy   3301 CACGCTGTCAGCGAGTGGTCATGACTGGTATAGACCCCCGGAACTTGGACTGGAAC 3360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 CACGCTGTCAGCGAGTGGTCATGACTGGTATAGACCCCCGGAACTTGGACTGGAAC 3360

Qy   3361 TATCTTCAGATTAGGATCTCTCGGCCACTAGCTACATGGGATGTTGTTTGTGACAACGTGTT 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 TATCTTCAGATTAGGATCTCTCGGCCACTAGCTACATGGGATGTTGTTTGTGACAACGTGTT 3420

Qy   3421 GTATACTGCTCACCATGGCAGCAAGGGCGCGCCCGGTTGGCTCATCCCAACCTCTATACA 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 GTATACTGCTCACCATGGCAGCAAGGGCGCGCCCGGTTGGCTCATCCCAACCTCCATACA 3480

```
Qy  3481  CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG  3540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481  CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG  3540

Qy  3541  GTCCCTTACTCGGTGCTCTTGCGGGAGACCAAGGGTATCTGTAACACGACTGGGGTC      3600
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541  GTCCCTTACTCGGTGCTCTTGCGGGAGACCAAGGGTATCTGTAACACGACTGGGGTC      3600

Qy  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGCGGGGCCCTTCCCAT     3660
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGCGGGGCCCTTCCCAT     3660

Qy  3661  GGCTGTTGCCAAGGGTTCTTCAGGTGCCCCGATTCTGTGCTCCTCCCGGGCATGTTATTGG  3720
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661  GGCTGTTGCCAAGGGTTCTTCAGGTGCCCCGATTCTGTGCTCCTCCCGGGCATGTTATTGG  3720

Qy  3721  GATGTTCACCGGCTGCTGTAGAAATTCTGGCCGTTCAGTCAGTCAGATTAGGGTTAGGCCGTT  3780
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721  GATGTTCACCGGCTGCTGTAGAAATTCTGGCCGTTCAGTCAGTCAGATTAGGGTTAGGCCGTT  3780

Qy  3781  GGTGTGTCTGGGATACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC    3840
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781  GGTGTGTCTGGGATACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC    3840

Qy  3841  TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC    3900
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841  TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC    3900

Qy  3901  CAAATTACCACTTCTTACACGAGGAGTATGAGGTCTTGGTCCTAAATCCCAGTGT      3960
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901  CAAATTACCACTTCTTACACGAGGAGTATGAGGTCTTGGTCCTAAATCCCAGTGT      3960

Qy  3961  GGCTACACAGCATCAAATGCCAAAGTGTACCAGGGGCTTCACTTACGTACGCGACGTACGGCGACAAATTG  4020
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961  GGCTACACAGCATCAAATGCCAAAGTGTACCAGGGGCTTCACTTACGTACGCGACGTACGGCGACAAATTG  4020

Qy  4021  CTATTTTAATGCCAACACAGGGGCTTCACTTACGTACAGCACATATGGCAT       4080
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021  CTATTTTAATGGCAAATGTACCAACACAGGGGCTTCACTTACGTACCACACATGGCAT       4080

Qy  4081  GTACCTGACCGGAGCATGTTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGC     4140
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081  GTACCTGACCGGAGCATGTTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGC     4140

Qy  4141  TACCGATGCAACCACCGTGTTGGGCATTGGGAAAGTCCTAACCGAAGCTCCATCCAAAAA   4200
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141  TACCGATGCAACCACCGTGTTGGGCATTGGGAAAGTTCCTAACCGAAGCTCCATCCAAAAA   4200

Qy  4201  TGGTTAGGCTAGTTGCTTGCTAACGAGTTAGCTCGAAGGAATAACAGCTGTCTCTTCTTAGGGG  4260
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201  TGGTTAGGCTAGTTGCTTGCTAACGAGTTAGCTCGAAGGAATAACAGCTGTCTCTTCTTAGGGG  4260

Qy  4261  CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAGAT    4320
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261  CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAGAT    4320

Qy  4321  TAAGGAGGAAAATCTGAAGAAAGGAGACACCTTATCTTTGAGGGCTACCAAAAACACTG    4380
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321  TAAGGAGGAAAATCTGAAGAAAGGAGACACCTTATCTTTGAGGGCTACCAAAAACACTG    4380

Qy  4381  TGATGAGCTTGCTAACGAGTTAGCTCGAAGGAATAACAGCTGTCTCTTCTTACTATAGGGG  4440
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381  TGATGAGCTTGCTAACGAGTTAGCTCGAAGGAATAACAGCTGTCTCTTCTTACTATAGGGG  4440

Qy  4441  ATGTGACATCTCAAAAATCTCAAAAATGTTGTAGTTGCCACTGATGCCTTGTG       4500
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441  ATGTGACATCTCAAAAATCTCAAAAATGTTGTAGTTGCCACTGATGCCTTGTG       4500

Qy  4501  TACAGGGTACACTCGGTGACTTTGATTCCGTGATGACTGACTTGTGTAGTGCGAAGGCAC    4560
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501  TACAGGGTACACTCGGTGACTTTGATTCCGTGATGACTGACTTGTGTAGTGAGAAGGCAC    4560

Qy  4561  ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGGTTCGTGTGTGCGGGGTTTCAGC   4620
```

```
Db  4561  ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGGTTCGTGTGTGCGGGGTTTCAGC   4620

Qy  4621  AATAGTTAAAGGCCAGCGGTAGGGCCGTAGGGGAGAGCTGGCATATACTACTA        4680
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621  AATAGTTAAAGGCCAGCGGTAGGGCCGTAGGGGAGAGCTGGCATATACTACTA        4680

Qy  4681  TGTAGACGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT     4740
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681  TGTAGACGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT     4740

Qy  4741  CGACGCAGCCAAGGCATGTATGGTTGTGTCATCAACAGAGAGTCAAACTATTCTGGACAC   4800
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741  CGACGCAGCCAAGGCATGTATGGTTGTGTCATCAACAGAGAGTCAAACTATTCTGGACAC   4800

Qy  4801  CTATCGCACCACCAACCTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT   4860
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801  CTATCGCACCACCAACCTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT   4860

Qy  4861  CTTTTCTATGGTCAACCCGAACCTTCATTTGTCAATACTGTCAAAAGAACTGCTGACAA    4920
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861  CTTTTCTATGGTCAACCCGAACCTTCATTTGTCAATACTGTCAAAAGAACTGCTGACAA    4920

Qy  4921  TTATGTTTGTTGACTGGCAGCCCCAACTACAACTGTCATCAGTATGGCTATGCTGCTCC    4980
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921  TTATGTTTGTTGACTGGCAGCCCCAACTACAACTGTCATCAGTATGGCTATGCTGCTCC    4980

Qy  4981  CAATGACGCACCACGGTGGCGGCTGACGCGCTGCCTGCCTGCGCCCAGAGCCCAGCCCAGCCCGGCTCCGGTTCTGTG  5040
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981  CAATGACGCACCACGGTGGCGGCTGACGCGCTGCCTGCCTGCGCCCAGAGCCCAGCCCAGCCCGGCTCCGGTTCTGTG  5040

Qy  5041  GCGCTTGACGCGGCTGACGCGCTGCTGCCTGCCTGCGCCCAGAGCCCAGCCCAGCCCGGTTGACGAGGTGACCAGATACCA  5100
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041  GCGCTTGACGCGGCTGACGCGCTGCTGCCTGCCTGCGCCCAGAGCCCAGCCCAGCCCGGTTGACGAGGTGACCAGATACCA  5100

Qy  5101  AATGTGCTTCACTGAAGTCAATACTTCACTTCTTGGACACAGCAGCCCGGCTGTGTCGGCCGTTGCGGTTCTTGCTGGAGT  5160
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101  AATGTGCTTCACTGAAGTCAATACTTCACTTCTTGGACACAGCAGCCCGGCTGTGTCGGCCGTTGCGGTTCTTGCTGGAGT  5160

Qy  5161  GGCTATGGCTATCTAGCCTACCCGTGTACTGTCCGCCTGCGGTTGCGGTTGCGGTTGCTTGCTGGTC  5220
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161  GGCTATGGCTATCTAGCCTACCCGTGTACTGTCCGCCTGCGGTTGCGGTTGCGGTTGCTTGCTGGTC  5220

Qy  5221  TATTACATCAGTCCCTACCCGGTGCACGCGTTGACGTGTCGCCCCAGTGGTTGACGAAGAAATCGT    5280
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221  TATTACATCAGTCCCTACCCGGTGCACGCGTTGACGTGTCGCCCCAGTGGTTGACGAAGAAATCGT    5280

Qy  5281  GGAGGAGTTGCATCAATCCCTTGCATCATTCATTCCCGAGGCCCATGTGCTGCAATGTGCTGATAAGCTGAA   5340
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281  GGAGGAGTTGCATCAATCCCTTGCATCATTCATTCCCGAGGCCCATGTGCTGCAATGTGCTGATAAGCTGAA   5340

Qy  5341  GAGTACAATCACCAACTAGTCTTCCTTCACATTGGAAAACCGCCCCTGAAAAACTTAACAC  5400
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341  GAGTACAATCACCAACTAGTCTTCCTTCACATTGGAAAACCGCCCCTGAAAAACTTAACAC  5400

Qy  5401  CTTTTCTTGGGCCTCATGGACCTACAATCCTTGCTACTCGTTGCTTTCATTGCGGGTTGGTTTTAGT    5460
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401  CTTTTCTTGGGCCTCATGGACCTACAATCCTTGCTACTCGTTGCTTTCATTGCGGGTTAGT    5460

Qy  5461  CACTTTACCTGACAATCCCTTTGCATCATGCGTGTTCATTGCGAGGCGCAATTGCGTCCAA    5520
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461  CACTTTACCTGACAATCCCTTTGCATCATGCGTGTTCATTGCGAGGCGCAATTGCGTCCAA    5520

Qy  5521  CCCACTACCTCACAAGATCAAAATGTTCCGTGTCATTATTTGGAGGCGCAATTGCCTCCAA    5580
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521  CCCACTACCTCACAAGATCAAAATGTTCCGTGTCATTATTTGGAGGCGCAATTGCCTCCAA    5580

Qy  5581  GCTTACAGACGCTAGAGGCGCACTGGCGTTCATGATGGCCGGGCTGCGCGGGAACAGCTCT    5640
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581  GCTTACAGACGCTAGAGGCGCACTGGCGTTCATGATGGCCGGGCTGCGCGGGAACAGCTCT    5640

Qy  5641  TGGTACAATGGACAATCGGTGGGTTTGTTCTTTGACATGGGTTCGTGTGTGCGGGGTTTCAGC   5700
```

Qy 6781 AGCTGGCGTTGACACCACCAAACTGCCAGCCCCTCCATCGAAGAGTAGTGGTAAGAAA 6840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6781 AGCTGGCGTTGACACCACCAAACTGCCAGCCCCTCCATCGAAGAGTAGTGGTAAGAAA 6840

Qy 6841 GGGCCAGTCCGGCAAGAACTGGTTGCGCTTGCCTCCCCCCTCCGAGATCCGTCCC 6900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6841 GGGCCAGTCCGGCAAGAACTGGTTGCGCTTGCCTCCCCCCTCCGAGATCCGTCCC 6900

Qy 6901 AGGAGTGTCATGTCCTGAAAGCTGCAACGAAGTGACCCGGTTAGAAGGTCCTTCAAACCT 6960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6901 AGGAGTGTCATGTCCTGAAAGCTGCAACGAAGTGACCCGGTTAGAAGGTCCTTCAAACCT 6960

Qy 6961 CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGCCATGCCGGTGGGAGCGGGTGA 7020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6961 CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGCCATGCCGGTGGGAGCAGGTGA 7020

Qy 7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGGCGAACAGGCGGAGGGCCCTGATGA 7080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGGCGAACAGGCGGAGGGCCCTGATGA 7080

Qy 7081 TTTACCCAGTTACCCTCCCAAAAAGAGAGTCTCTGAATGGTCAGACGAAAGTTGGTCGAC 7140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7081 TTTACCCAGTTACCCTCCCAAAAAGAGAGTCTCTGAATGGTCAGACGAAGTTGGTCAAC 7140

Qy 7141 GGCTACAACCGCTTCAGCTACGTACTGGCCCCCCGTACCCTAAGGACGGGGAAAGGA 7200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7141 GACTACAACCGCTTCAGCTACGTACTGGCCCCCCGTACCCTAAGGACGGGGAAAGGA 7200

Qy 7201 TTCCACTCAGTCAGCCCCGCCAAACCGCCTACAAAAAGAAGTTGGGAAAGAGTGAGTT 7260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7201 TTCGTCTCAGTCAGCCCCGCCAAACCGCCTACAAAAAGAAGTTGGGAAAGAGTGAGTT 7260

Qy 7261 TTCGTGCAGCATGAGCTGACCGAGCTGATTAGCTTCAAAACTGCTTCTAAAGT 7320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7261 TTCGTGCAGCATGAGCTGACCGAGCTGATTAGCTTCAAAACTGCTTCTAAAGT 7320

Qy 7321 TCTGTCTGCAACTCGGGCACTAGTGGTTTCCTCAAACAAGATCATTGGTATGT 7380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7321 TCTGTCTGCAACTCGGGCACTAGTGGTTTCCTCAAACAAGATCATTGGTATGT 7380

Qy 7381 GACTGAGCGCGGATGCGGCGGAGCTTAGAAACAAAAGTCACTATTAATAGACAACCTCT 7440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7381 GACTGAGCGCGGATGCGGCGGAGCTTAGAAACAAAAGTCACTATTAATAGACAACCTCT 7440

Qy 7441 GTTCCCCCATCATACACAACAGCAAGTGAGATTGGCTGAAGGAAAAGCTTCAAAAGTTGT 7500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7441 GTTCCCCCATCATACACAACAGCAAGTGAGATTGGCTGAAGGAAAAGCTTCAAAAGTTGT 7500

Qy 7501 CGGTGTCATGTGGGGCACTCGGGGCACTGAGCTGTCGTTCGTTCTGGAGCGCCCGTGTTCT 7560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7501 CGGTGTCATGTGGGGCACTCGGGGCACTGAGCTGTCGTTCGTTCTGGAGCGCCCGTGTTCT 7560

Qy 7561 CCACATCACTGGCCTTCGGGGCGTGATGTTCAGTTGTTATCTGGAGGCCCGCAAGGCCCAA 7620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7561 CCACATCACTGGCCTTCGGGGCGTGATGTTCAGTTGTTATCTGGAGGCCCGCAAGGCCCAA 7620

Qy 7621 GGACTTGCAGAAGTGTGTTGCAGGCAGGTGAGATACCGGACTCATTATCGGCAAACTGTGAT 7680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7621 GGACTTGCAGAAGTGTGTTGCAGGCAGGTGAGATACCGGACTCATTATCGGCAAACTGTGAT 7680

Qy 7681 AGTTCCAAAGGAGGAGGTCTTCGTGAGAGACCCCCAGAAACCAACAAAGAAACCCCCAAG 7740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAAACCAACAAGAAACCCCCAAG 7740

Qy 7741 GCTTATCTCGTACCCACCTTGAAATGAGAGATGTGTTGAGAGATGTTACGGTCAGGT 7800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7741 GCTTATCTCGTACCCACCTTGAAATGAGAGATGTGTTGAGAGATGTTACGGTCAGGT 7800

Qy 7801 TGCTCCTGACGTAGTTGATGCTGTAAAGCTGTCATGGGAGATCGTACGGGTTTGTGAGATCCACGTAC 7860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7801 TGCTCCTGACGTAGTTGATGCTGTAAAGCTGTCATGGGAGATCGTACGGGTTTGTGACCCCACGTAC 7860

Db 5641 TGGTACATGGACATCGGTGGGTTTGTCTTTGACATGCTAGGCGGCTATGCTGCCGCCTC 5700

Qy 5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGGTGGCCCACTATGGATCAGCT 5760
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGGTGGCCCACTATGGATCAGCT 5760

Qy 5761 TGCTGGTTAGTCTACTCCCGCGTTCAATCCGGCCCGCCGAGGAGTTGTGGCGTCTTGTCAGC 5820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5761 TGCTGGTTTAGTCTACTCCGCGTTCAATCCGGCCCGCCGAGGAGTGTGGCGTCTTGTCAGC 5820

Qy 5821 TTGTGCAAATGTTTGCTTTGACAACAGCAGGGCCAGATCACTGGCCAACAGACTTCTTAC 5880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5821 TTGTGCAAATGTTTGCTTTGACAACAGCAGGGCCAGATCACTGGCCAACAGACTTCTTAC 5880

Qy 5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940

Qy 5941 CAGGAAGATACTGGGCATTCGGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT 6000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5941 CAGGAAGATACTGGGCATTCGGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT 6000

Qy 6001 CCGTTGGCTCCACACCCCGACGGAGGAGGATGATTGCGGCCTCATTGGGTCTCAGAGAT 6060
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6001 CCGTTGGCTCCACACCCCGACGGAGGAGGATGATTGCGGCCTCATTGGGTCTCAGAGAT 6060

Qy 6061 TTGGCAGTATGTGCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAGCTGGAGTTCA 6120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6061 TTGGCAGTATGTGCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAGCTGGAGTTCA 6120

Qy 6121 GAGCATGGTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAGGGGGTACAAGGGCCC 6180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6121 GAGCATGGTTACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAGGGGGTACAAGGGCCC 6180

Qy 6181 CTGGATTGGATCAGGTAATGCTCAGGCACGCGTGTCCAAGCACGCGTGTCGAACTCATCTTTTC 6240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6181 CTGGATTGGATCAGGTAATGCTCAGGCACGCGTGTCCAAGCACGCGTGTCGAACTCATCTTTC 6240

Qy 6241 TGTTGAGAATGTTTGTCAAAATTACAAGAGACCCCAGAACTTGTTCAAATTACTACTGGAG 6300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6241 TGTTGAGAATGTTTGTCAAAATTACAAGGACCCCAGAACTTGTTCAAATTACTACTGGAG 6300

Qy 6301 AGGGGCTGTTCCAGTCAACGCTCAGGCGTGTGTGGGTGTTCCGGCTGTGTGGGTGTCCGGCTGTGTGGGTGTGATTG 6360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6301 AGGGGCTGTTCCAGTCAACGCTCAGGCGTGTGTGGGTGTGATTG 6360

Qy 6361 GACTAGTCTTGTCGTCAATTATGGCGTAGGACTACTGTGTAAAATGAGAAATGGGAGA 6420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6361 CTTAGTCTTGTCGTCAATTATGGCGTAGGACTACTGTGTAAAATGAGAAATTGGGAGA 6420

Qy 6421 TCACATTTTGTTACAGCAGACATCTGCTTGTTCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCCAAC 6480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6421 TCACATTTTGTTACAGCAGACATCTGCTTGTTCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCCAAC 6480

Qy 6481 CTTGAGAGACTGCAGTGGCCGTGGACGGCGTGGACGGCGTACAGGTTCAGTTGTTATCTGGTGTTGAGCCCAA 6540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6481 CTTGAGAGACTGCAGTGGCCGTGGACGGCGTACAGGTTCAGTTGTTATCTGGTGTTGAGCCCAA 6540

Qy 6541 AACTCCTTGGACGACATCTGCTTGCTTGCCGTGGTCGTGCTACGGTGTAAAACTGTTAA 6600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6541 AACTCCTTGGACGACATCTGCTTGCTTGCCGTGGTCGTGCTACGGTGTAAAACTGTTAA 6600

Qy 6601 GCTTCCCTTCCGGGTTGACGGTGACGGTCACACAACACCTGGTGTGCGCATGCAACTTAATTTGCGTGA 6660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6601 GCTTCCCTTCCGGGTTGACGGTCACACAACACCTGGTGTGCGCATGCAACTTAATTTGCGTGA 6660

Qy 6661 TGCACTTGAGACAAATGCAGTGCAGTTGCCGGCGTACACTCCTAGTGATGAAGCCCCAGT 6720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6661 TGCACTTGAGACAAATGCAGTGCAGTTGCCGGCGTACACTCCTAGTGATGAAGCCCAGT 6720

Qy 6721 GTCCGCTCTTGTTTCAAACAGGAGTTGCCGGCGTACAAACCCAATTGCTTGAGGCAATTTC 6780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6721 GTCCGCTCTTGTTTCAAACAGGAGTTGCCGGCGTACAAACCCAATTGCTTGAGGCAATTTC 6780

```
Qy  7861 CCGTGTCAAGCGTCGTGTGTCGATGTGGTCAGTCGAGCCACATGCGATAC 7920
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7861 CCGTGTCAAGCGTCGTGTGTCGATGTGGTCAGTCGAGCCACATGCGATAC 7920

Qy  7921 AGTGTGTTTGACAGTACCATCACACCCGAGATATCATGGTGGAGACAGACATCTACTC 7980
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7921 AGTGTGTTTGACAGTACCATCACACCCGAGATATCATGGTGGAGACAGACATCTACTC 7980

Qy  7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGCCATTCACACCATTGCGAGGCAGTTATA 8040
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGCCATTCACACCATTGCGAGGCAGTTATA 8040

Qy  8041 CGCTGGAGGACCCATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGTGTAGGTC 8100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8041 CGCTGGAGGACCCATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGTGTAGGTC 8100

Qy  8101 TTCCGGGGTCTATACTCCAAGTTCCAACAGTTTGACCTGCTGGCTGAAGGTAAATGC 8160
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8101 TTCCGGGGTCTATACTCCAAGTTCCAACAGTTTGACCTGCTGGCTGAAGGTAAATGC 8160

Qy  8161 TGCAGCCGAACAGGCTGCATGAAGAACCCTCGCTTCCTTATTTGCGGCGATGATTGCAC 8220
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161 TGCAGCCGAACAGGCTGCATGAAGAACCCTCGCTTCCTTATTTGCGGCGATGATTGCAC 8220

Qy  8221 CGTAATTTGGAAGAGCGCCGGAGCAGATGCAGACAAAGCAATGCGTGTCTTGCTAG 8280
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8221 CGTAATTTGGAAGAGCGCCGGAGCAGATGCAGACAAAGCAATGCGTGTCTTGCTAG 8280

Qy  8281 CTGGATGAAGGTGATGGGTGACCACCAAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281 CTGGATGAAGGTGATGGGTGACCACCAAGATTGTGTGCCTCAACCCAAATACAGTTTGA 8340

Qy  8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400

Qy  8401 CTACTTTCTTACAAGATCCTCGTATCCCCCTTGGCAGTGCTCTGCCGAGGGTCTCGGG 8460
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8401 CTACTTTCTTACAAGATCCTCGTATCCCCCTTGGCAGTGCTCTGCCGAGGGTCTCGGG 8460

Qy  8461 ATACAACCCCAGTGCTGCGTGGATTGGTATCTAATACATCACTACCCATGTTGTGGGT 8520
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8461 ATACAACCCCAGTGCTGCGTGGATTGGTATCTAATACATCACTACCCATGTTGTGGGT 8520

Qy  8521 TAGCCGTGTGTGTGGCTGTCCATTTCATGTGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8521 TAGCCGTGTGTGTGGCTGTCCATTTCATGTGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580

Qy  8581 GACTGTGACCTTTGACTGTGTATGGGAAAAATTATACGGTGCCGTGTAGAAGATCTGCCCAG 8640
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8581 GACTGTGACCTTTGACTGTGTATGGGAAAAATTATACGGTGCCGTGTAGAAGATCTGCCCAG 8640

Qy  8641 CATCATTGCTGGGTGTGCACGGTATTGAGGCTTTCTCCGTGGTGCGCTACCAACGCTGA 8700
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8641 CATCATTGCTGGGTGTGCACGGTATTGAGGCTTTCTCCGTGGTGCGCTACCAACGCTGA 8700

Qy  8701 GATCCTCAGAGTGCTGCGTGGATTGGTATCCAATCACTACAAGACATGACCATGGGCGCG 8760
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8701 GATCCTCAGAGTGCTGCGTGGATTGGTATCCAATCACTACAAGACATGACCATGGGCGCG 8760

Qy  8761 AAAGAAAGCCAGGGACGTCCTGCCAGCGCGTGGCGAGCACACCCAAATT 8820
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8761 AAAGAAAGCCAGGGACGTCCTGCCAGCGCGTGGCGAGCACACCCAAATT 8820

Qy  8821 GGCTCGCTTCCTTCTCTGGCATGCTACATCTACCAGATTTTGGATAAGACGAG 8880
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8821 GGCTCGCTTCCTTCTCTGGCATGCTACATCTACCAGATTTTGGATAAGACCAG 8880

Qy  8881 CGTGGCTCGGTACACCACTTCAATTATTGTGATGTTTACTCGCCGAGGGGATGTGTT 8940
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8881 CGTGGCTCGGTACACCACTTCAATTATTGTGATGTTTACTCGCCGAGGGGATGTGTT 8940

Qy  8941 TATTACACCACGAGAGAAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTGTCATTGTTTT 9000
```

```
Db  8941 TGTTACACCACCAGAAGACGAAGATTGCAGAGAAGTTTCTTGTGAAGTATTTGGCTGTGTCATTGTTTT 9000
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  9001 TGCCCTAGGGCTCATTGCTGTTGGACTGGTTGGACTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9001 TGCCCTAGGGCTCATTGCTGTTGGACTGGTTGGACTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060

Qy  9061 CTAACAG----TTTTTTTTTTTTTTTTTTT 9116
         ||||||   ||||||||||||||||||
Db  9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGCAACAGGGGAGACCCC 9120

Qy  9117 GGGCTTAACGACCCCGC 9133
         |||||||||||||||||
Db  9121 GGGCTTAACGACCCCGC 9137
```

RESULT 14
AR494272

LOCUS   AR494272   9143 bp   DNA   linear   PAT 15-MAY-2004
DEFINITION Sequence 390 from patent US 6720166.
ACCESSION   AR494272
VERSION   AR494272.1 GI:47267258
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 9143)
  AUTHORS   Simons,J.N., Pilot-Matias,T.J., Dawson,G.J., Schlauder,G.G.,
            Desai,S.M., Leary,T.P., Muerhoff,A.S., Erker,J.C., Buijk,S.L. and
            Mushahwar,I.K.
  TITLE   Non-a, non-b, non-c, non-d, non-e hepatitis reagents and
            methods for their use
  JOURNAL   Patent: US 6720166-A 390 13-APR-2004;
FEATURES   Location/Qualifiers
     source   1..9143
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match   96.4%;   Score 9059.8;   DB 6;   Length 9143;
Best Local Similarity 99.6%;   Pred. No. 0;
Matches 9096;   Conservative 0;   Mismatches 37;   Indels 4;   Gaps 1;

```
Qy    1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCTAG   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCTAG   60

Qy   61 CAGGGCGTGGGGGGATTTCCCCGTGCGAGAAGAGGTGGAGCCAACCACCACCTTAGTAT  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CAGGGCGTGGGGGGATTTCCCCGTGCGAGAAGAGGTGGAGCCAACCACCACCTTAGTAT  120

Qy  121 GTAGGCGCGGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTAGGCGCGGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180

Qy  181 CCTTGATGGGCGTTCATGGGTTCGGTGGTGGTGGTAGGGGAAGACCGGGGACCGGGGACACG  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CCTTGATGGGCGTTCATGGGTTCGGTGGTGGTGGTAGGGGAAGACCGGGGACCGGGGACACG  240

Qy  241 CCTCCCAGATAGAGCGCGGCGGCACTGTAGGAGAAGACCGGGGACCGGGGACTACCAAGGACG  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 CCTCCCAGATAGAGCGCGGCGGCACTGTAGGAGAAGACCGGGGACCGGGGACTACCAAGGACG  300

Qy  301 CAGACCTCTTTTGAGTATCACGCTCGGGAAGTAGTTGGGGCAAGCCCACCTATATGTGT  360
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGACCTCTTTTGAGTATCACGCTCGGGAAGTAGTTGGGGCAAGCCCACCTATATGTGT  360

Qy  361 TGGGATGGTTGGGGTTAGCCATCCATCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 TGGGATGGTTGGGGTTAGCCATCCATCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT  420

Qy  421 CTGGGAGTCCGTAGACCGTAGCACATGCGTGTATTCTACTCAAACAAGTCCTGTACC  480
```

Qy 421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTGTTATTTCTACTCAAACAAGTCCTGTACC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 TGCGCCCAGAACGCGCAAGAACAAGCAGACGCGAGGCTTCATATCCTGTCCATTAAAAC 540

Qy 481 TGCGCCCAGAACGCGCAAGAACAAGCAGACGCGAGGCTTCATATCCTGTCCATTAAAAC 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 ATCTGTTGAAAGGGACAACGAGCAAAGTCCAGCGGATGCTCAGGCTCCGGCCCTCGTAA 600

Qy 541 ATCTGTTGAAAGGGACAACGAGCAAAGTCCAGCGGATGCTCAGGCTCCGGCCCTCGTAA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 TTACAAAATTGCTGGTATCCATGATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC 660

Qy 601 TTACAAAATTGCTGGTATCCATGATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 TCATGGTTGGGGACGCCAAGACCCTGCCAATCTTGGAATCCTTCTGA 720

Qy 661 TCATGGTTGGGGACGCCAAGACCCTGCCAATCTTGGAATCCTTCTGA 720
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 TTACCCTTTGGGGTGATTGGTGATGTTACAACTCACACACCTCTAGTAGGCCCGCTGGT 780

Qy 721 TTACCCTTTGGGGTGATTGGTGATGTTACAACTCACACACCTCTAGTAGGCCCGCTGGT 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 781 GGCAGGAGCGGGTCGTTCGACCAGTCTGCCAGATAGTACGCTTGGAGGATGGAGTCAA 840

Qy 781 GGCAGGAGCGGGTCGTTCGACCAGTCTGCCAGATAGTACGCTTGGAGGATGGAGTCAA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 CTGGGCTACTCGTTGGTTCGGTGTCCACCTTTTGTGGTATGTCTGTATCTTTGCCCTG 900

Qy 841 CTGGGCTACTCGTTGGTTCGGTGTCCACCTTTTGTGGTATGTCTGTATCTTTGCCCTG 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 TCCCTGTAGTTGGGGGCGGGTCACTGACCCAGACACAAATACCACAATCCTGACCAATTG 960

Qy 901 TCCCTGTAGTTGGGGGCGGGTCACTGACCCAGACACAAATACCACAATCCTGACCAATTG 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 CTGCCAGCGTAATCAGGTTATCTATTGTTCCTTCCACTTGCCTACACGAGCCTGGTTG 1020

Qy 961 CTGCCAGCGTAATCAGGTTATCTATTGTTCCTTCCACTTGCCTACACGAGCCTGGTTG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 TGTGATCTGCGCGGACGAGTGCTGGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA 1080

Qy 1021 TGTGATCTGCGCGGACGAGTGCTGGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 TTGGACTGGCACGGACTCCTTCTTGGCTGACCACATTGTTGTGTGTATTGGTGGGCTCTTGT 1140

Qy 1081 TTGGACTGGCACGGACTCCTTCTTGGCTGACCACATTGTTGTGTGTATTGGTGGGCTCTTGT 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 GACCTGTCAGGGTATTGGTGGTTTGTGTCGTGTGTTGTGTGTTAGTCGGTGACTG 1200

Qy 1141 GACCTGTCAGGGTATTGGTGGTTTGTGTCGTGTGTTGTGTGTTAGTCGGTGACTG 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 GCTTGTCAGGCACTGGCTTATTCACATAGAACTGGTACTTGTTACCTGGA 1260

Qy 1201 GCTTGTCAGGCACTGGCTTATTCACATAGAACTGGTACTTGTTACCTGGA 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGGATGGCCAAGGT 1320

Qy 1261 AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGGATGGCCAAGGT 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT 1380

Qy 1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 GTTTAGCAGTGTACACTACCTGGCCGGTTGGCGGTTGGGCGTCTGATCTACTATGCCTCTCGGGGCAA 1440

Qy 1381 GTTTAGCAGTGTACACTACCTGGCCGGTTGGCGGTTGGGCGTCTGATCTACTATGCCTCTCGGGGCAA 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 GTGGTATCAGTTGCTCTCCTAGCGCTTATGCTTTACATAGAAGCGACCTCTGGAAACCCCAT 1500

Qy 1441 GTGGTATCAGTTGCTCTCCTAGCGCTTATGCTTTACATAGAAGCGACCTCTGGAAACCCCAT 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTTGCTCGCCTTTGATGATACCATGTCC 1560

Qy 1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTTGCTCGCCTTTGATGATACCATGTCC 1560

Db 1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTGTACAGTCCAAAGTGGAC 1620

Qy 1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTGTACAGTCCAAAGTGGAC 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 CAGGCCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCTATACAATCCCTGG 1680

Qy 1621 CAGGCCCTGTCACTCTTCAGAGTATAACAACTCCATATCTTGGTACCCTATACAATCCCTGG 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 TGCGAGGGGATGTATGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA 1740

Qy 1681 TGCGAGGGGATGTATGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 TGTGCCATCGTACTGCGACTATGGCACTGATGCAGTGCAGTGTGGAACGACACTCGCAACACTTA 1800

Qy 1741 TGTGCCATCGTACTGCGACTATGGCACTGATGCAGTGCAGTGTGGAACGACACTCGCAACACTTA 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 CGAAGCATGCGGTGTAACACCATGGCTAACAACCGCATGGCACAACGGCTCAGCCCTGAA 1860

Qy 1801 CGAAGCATGCGGTGTAACACCATGGCTAACAACCGCATGGCACAACGGCTCAGCCCTGAA 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAAATGTTTAAACCTCATAATTGGATGTC 1920

Qy 1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAAATGTTTAAACCTCATAATTGGATGTC 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921 AGGCCATTGTATTTGAGGGATCAGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA 1980

Qy 1921 AGGCCATTGTATTTGAGGGATCAGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981 TTTCCACTCTCCTACCGGAGAGGTGGGTAGGTTGCCCGGTACCCCACCTGTGGTACG 2040

Qy 1981 TTTCCACTCTCCTACCGGAGAGGTGGGTAGGTTGCCCGGTACCCCACCTGTGGTACG 2040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041 TGGTTCTGGTTACAGTTTCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100

Qy 2041 TGGTTCTGGTTACAGTTTCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101 ATTGATACCACAAGACAAAGCCTGGAAAAATTATCAGTCTTATATTCCGCCACGGGTGC 2160

Qy 2101 ATTGATACCACAAGACAAAGCCTGGAAAAATTATCAGTCTTATATTCCGCCACGGGTGC 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161 TTTGTCTCTTACGGGAGTTACCACCAGGCCGTGGTGTCTAATTCTGTTGGGGTTGTGTGG 2220

Qy 2161 TTTGTCTCTTACGGGAGTTACCACCAGGCCGTGGTGTCTAATTCTGTTGGGGTTGTGTGG 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221 CAGCAAGTATCTTATTTAGCCTACCTTTGCTTTTGATTTCTTCCTCTGTCTGTTGGGGCGCTTC 2280

Qy 2221 CAGCAAGTATCTTATTTAGCCTACCTTTGCTTTTGATTTCTTCCTCTGTCTGTTGGGGCGCTTC 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2281 TGGTTACCCTTTGCGTCCGTGTCGTATCGTCAAGCTGGCTGGGATGT 2340

Qy 2281 TGGTTACCCTTTGCGTCCGTGTCGTATCGTCAAGCTGGCTGGGATGT 2340
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 2341 TTTGTCTAAAGCTCAAGTAGCTCCTTTTGCTTTCCTCGTGTTGCTATCTCCG 2400

Qy 2341 TTTGTCTAAAGCTCAAGTAGCTCCTTTTGCTTTCCTCGTGTTGCTATCTCCG 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401 CTGCAGGCTACGTTGTGTCCGCGCGGTTGTGTGCCCATGGCTGCGGGGCTTC 2460

Qy 2401 CTGCAGGCTACGTTGTGTCCGCGCGGTTGTGTGCCCATGGCTGCGGGGCTTC 2460
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2461 AACTTTCTTTGTTGCAGCAGTTATGACTGGTCGGTGCGGACTGCT 2520

Qy 2461 AACTTTCTTTGTTGCAGCAGTTATGACTGGTCGGTGCGGACTGCT 2520
         |||||||||||||||||||||||||||||||||||||||||||||
Db 2521 AGTGGGCAGGGTTAGTTTTGTGTGGGCCGTAACCGTGGTGCCACTCTGCTTGT 2580

Qy 2521 AGTGGGCAGGGTTAGTTTTGTGTGGGCCGTAACCGTGGTGCCACTCTGCTTGT 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2581 AGTTCCTTGGCCTCTGGTAGCCGCTTTTAACCCTCTTGCATTTGGTTGGCTACGGCCTTCAGC 2640

Qy 2581 AGTTCCTTGGCCTCTGGTAGCCGCTTTTAACCCTCTTGCATTTGGTTGGCTACGGCCTTCAGC 2640

The page consists of a two-column nucleotide sequence alignment between a query sequence (Qy) and a database sequence (Db), with numerical position coordinates flanking each line of sequence.

Right column (positions 5941–7020):

```
Qy  5941 CAGGAAGATACTGGGCATTCTTGGAGGCATCACCCCCTGGAGTGTCATCATCAGCTTGCAT 6000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941 CAGGAAGATACTGGGCATTCTTGGAGGCATCACCCCCTGGAGTGTCATCATCAGCTTGCAT 6000

Qy  6001 CCGTTGGCTCCACACCCGACGGAGGATGAATTGCGGCCTCATTGCTTGGGGTCTTAGAGAT 6060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001 CCGTTGGCTCCACACCCGACGGAGGATGAATTGCGGCCTCATTGCTTGGGGTCTTAGAGAT 6060

Qy  6061 TTGGCAGTATGTGTGCAATTTCTTTGTGATTTGTATTTGCTTTAATGTCCTTAAAGCTGGAGTTCA 6120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061 TTTGGCAGTATGTGTGCAATTTCTTTGTGATTTGTATTTGCTTTAATGTCCTTAAAGCTGGAGTTCA 6120

Qy  6121 GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGTTGCCAGAAGGGGTACAAGGGCCC 6180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121 GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGTTGCCAGAAGGGGTACAAGGGCCC 6180

Qy  6181 CTGGAATTGGATCAGGTATGCTCCAAGCACGTGTCCAAGCACGCGTCTCGAACTCATCTTTTC 6240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181 CTGGAATTGGATCAGGTATGCTCCAAGCACGTGTCCAAGCACGCGTCTCGAACTCATCTTTTC 6240

Qy  6241 TGTTGAGAATGGTTTTGCAAAATCTTACAAACGACCCAGAACTTGTTCAAATTACTGGAG 6300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241 TGTTGAGAATGGTTTTGCAAAATCTTACAAACGACCCAGAACTTGTTCAAATTACTGGAG 6300

Qy  6301 AGGGGCTGTTCCAGTCAACGCTGTGTGGGTCGGTCGGCTGTAGACCGACCCAACTGATTG 6360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301 AGGGGCTGTTCCAGTCAACGCTGTGTGGGTCGGTCGGCTGTAGACCGACCCAACTGATTG 6360

Qy  6361 GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAATGGGAGA 6420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361 GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAATGGGAGA 6420

Qy  6421 TCACATTTGTTACAGCAGTATCCTCCAAATGTCGTTTCACCCAGGTGCCCCCCAAC 6480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421 TCACATTTTTGTTACAGCAGTATCCTCCAAATGTCGTTTCACCCAGGTGCCCCCCAAC 6480

Qy  6481 CTTGAGAGCTGCAGTGGCCGTGACGGCGTAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481 CTTGAGAGCTGCAGTGGCCGTGACGGCGTAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540

Qy  6541 AACTCCTTGGACGACATCGTCTGCTTGTACGGTAAGGGTAAAACTGTTAA 6600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541 AACTCCTTGGACGACATCGTCTGCTTGTACGGTAAGGGTAAAACTGTTAA 6600

Qy  6601 GCTTCCCTTCGCGGTTGACGGTGCCACACACCTGGTGCGCATGCGCAACTTAATTGCGTGA 6660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601 GCTTCCCTTCGCGGTTGACGGTGCCACACACCTGGTGCGCATGCGCAACTTAATTGCGTGA 6660

Qy  6661 TGCACTTGAGACAAATGACTGTAATTCCACAAACAACACTCCTAGTGATGAAGCCGCAGT 6720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661 TGCACTTGAGACAAATGACTGTAATTCCATAAACAACACTCCTAGTGATGAAGCCGCAGT 6720

Qy  6721 GTCCGCTCTTGTTTCAAACAGGAGTTGCGGCTATGCTGCCCCTCATCGAAGAGTGTTAAGAAA 6780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721 GTCCGCTCTTGTTTCAAACAGGAGTTGCGGCTATGCTGCCCCTCATCGAAGAGTGTTAAGAAA 6780

Qy  6781 AGCTGGCGTTGACACCAAACTGCCAGCCCCCTCATCGAAGAGGTAGTGGTAAGAAA 6840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781 AGCTGGCGTTGACACCAAACTGCCAGCCCCCTCATCGAAGAGGTAGTGGTAAGAAA 6840

Qy  6841 GGGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCTCCGAGATCCGTCCC 6900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCTCCGAGATCCGTCCC 6900

Qy  6901 AGGAGTGTCATGTCCTGAAAGCCTGCGAACGAGTGACCCGTTAGAAGGTCCTTCAAACCT 6960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901 AGGAGTGTCATGTCCTGAAAGCCTGCGAACGAGTGACCCGTTAGAAGGTCCTTCAAACCT 6960

Qy  6961 CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGCCGATGCCCGTGTTGGGAGCGGGTGA 7020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961 CCCTTCTTCACCACCTGTTCTACAGTTGGCCATGCCGATGCCCGTGTTGGGAGCGGGTGA 7020
```

Left column (positions 4801–5940):

```
Db  4801 CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTGGACGAGTGGGCTGATCT 4860

Qy  4861 CTTTTCTATGGTCAACCCGACCTTCATTTGTCAATACTGCAAAAAGAACTGCTGACAA 4920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861 CTTTTCTATGGTCAACCCGACCTTCATTTGTCAATACTGCAAAAAGAACTGCTGACAA 4920

Qy  4921 TTATGTTTGTGACTGCAGCCCAACTACAACTGTGTCATCAGTATGCTATGCTGCTCC 4980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921 TTATGTTTGTGACTGCAGCCCAACTACAACTGTGTCATCAGTATGCTATGCTGCTCC 4980

Qy  4981 CAATGACGACCACCGTGGGCAGGGAGCCCGGCTTGTGTGGGGTCTGTG 5040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981 CAATGACGACCACCGTGGGCAGGGAGCCCGGCTTGTGTGGGGTTCTGTG 5040

Qy  5041 GCGCTTGACGCGGCGCTGACGCCTGTCCTCGGCCCAGAGCCCAGCGAGGTGACCAGATACCA 5100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041 GCGCTTGACGCGGCGCTGACGCCTGTCCTCGGCCCAGAGCCCAGCGAGGTGACCAGATACCA 5100

Qy  5101 AATGTGCTTCACTGAAGTCAATACTTCGGGACAGCGCGACTCGCTGCTGTTGGCGTTGGAGT 5160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101 AATGTGCTTCACTGAAGTCAATACTTCGGGACAGCGCGACTCGCTGCTGTTGGCGTTGGAGT 5160

Qy  5161 GGCTATGGCTTATCTAGCCATTGACACTTTGGCGGCCACTTGTGTCGTGCGGCGTTGCTGGTC 5220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161 GGCTATGGCTTATCTAGCCATTGACACTTTGGCGGCCACTTGTGTCGTGCGGCGTTGCTGGTC 5220

Qy  5221 TATTACATCAGTCCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAGAAATCGT 5280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221 TATTACATCAGTCCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAGAAATCGT 5280

Qy  5281 GGAGGAGTGTGCATCATTCAATCCTTGGAGGCCATGGTTGCTGCAATTGACAAGCTGAA 5340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281 GGAGGAGTGTGCATCATTCAATCCTTGGAGGCCATGGTTGCTGCAATTGACAAGCTGAA 5340

Qy  5341 GAGTACAAATCACCACACCAGTAGTTCCTTTCACATTGGAAAACCGCCCCTGAAAACTTAACAC 5400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341 GAGTACAATAACCACACCAGTAGTTCCTTTCACATTGGAAAACCGCCCCTGAAAACTTAACAC 5400

Qy  5401 CTTTCTTGGGCTCATGCAGTCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGGTTAGT 5460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401 CTTTCTTGGGCTCATGCAGTCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGGTTAGT 5460

Qy  5461 CACTTTACCTGACAATCCCTTTGCATCATGCGGTTGCTTTCATTGCGGGTATTACTAC 5520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461 CACTTTTACCTGACAATCCCTTTGCATCATGCGGTTGCTTTCATTGCGGGTATTACTAC 5520

Qy  5521 CCCACTACCTCACACAAGATCAAGATCAAGCAGGGCACTGGGTGGTTCATTGGGCGTCCAA 5580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521 CCCACTACCTCACACAAGATCAAGATCAAGCAGGGCACTGGGTGGTTCATTGGGCGTCCAA 5580

Qy  5581 GCTTACAGACGCTAGGAGCGCACTCGGGTTTCATGATGGCGCACTGCGGGAACAGCTCT 5640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581 GCTTACAGACGCTAGGAGCGCACTCGGGTTTCATGATGGCGCACTGCGGGAACAGCTCT 5640

Qy  5641 TGGTACATGGACATCGGTGGGTTTGTCTTTGACATGCTAGGCGGCTATGCTGCCCGCTC 5700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641 TGGTACATGGACATCGGTGGGTTTGTCTTTGACATGCTAGGCGGCTATGCTGCCCGCTC 5700

Qy  5701 ATCCACTGCTTGCTTGACATTAAATGCTTGACATGGTGAGTGGCCCACTATGGATCAGCT 5760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701 ATCCACTGCTTGCTTGACATTAAATGCTTGACATGGTGAGTGGCCCACTATGGATCAGCT 5760

Qy  5761 TGCTCGGTTTAGTCTACTCCGCGTTCAATCCGCGTTCAATCCCGCGTTGTGGGCGTCTTGTCAGC 5820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761 TGCTCGGTTTAGTCTACTCCGCGTTCAATCCGCGTTCAATCCCGCGTTGTGGGCGTCTTTGTCAGC 5820

Qy  5821 TTGTGCAAATGTCATGTCCTGAAAGCCTGGCCCAACACTCGGCCCCAACAGACTTCTTAC 5880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821 TTGTGCAATGTTTGCTTGCTTGACACACTGCCCAACAGACTTCTTAC 5880

Qy  5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940
```

```
Qy   7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAACAGGCGGAGGCCCTGATGA 7080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAACAGGCGGAGGCCCTGATGA 7080

Qy   7081 TTTACCCAGTTACCCTCCCAAAAAGGAGGTCTCGAATGGTCAGACGAAGTTGGTCGAC 7140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7081 TTTACCCAGTTACCCTCCCAAAAAGGAGGTCTCGAATGGTCAGACGAAGTTGGTCAAC 7140

Qy   7141 GGCTACAACCGCTTCCAGCTACGTTACTGGCCCCCGTACCTTAAGTACGGGGAAAGGA 7200
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   7141 GACTACAACCGCTTCCAGCTACGTTACTGGCCCCCGTACCTTAAGATACGGGGAAAGGA 7200

Qy   7201 TTCCACTCAGTCAGCCCCGCCAAACGGCCTACAAAAAAGAAGTTGGGAAGAGTGAGTT 7260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7201 TTCCACTCAGTCAGCCCCGCCAAACGGCCTACAAAAAAGAAGTTGGGAAGAGTGAGTT 7260

Qy   7261 TTCGTGCAGCATGAGCTACACCTGGACCGACGTGATTTAGCTTCAAAACTGCTTCTAAAGT 7320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7261 TTCGTGCAGCATGAGCTACACCTGGACCGACGTGATTTAGCTTCAAAACTGCTTCTAAAGT 7320

Qy   7321 TCTGTCTGCAACTCGGGCCATCACTGGTTCCTCAAACAAAGATCATTGGTTGTATGT 7380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7321 TCTGTCTGCAACTCGGGCCATCACTGGTTCCTCAAACAAAGATCATTGGTTGTATGT 7380

Qy   7381 GACTGAGCGCGGGATGCGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCT 7440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7381 GACTGAGCGCGGGATGCGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCT 7440

Qy   7441 GTTCCCCCATCATACCACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7441 GTTCCCCCATCATACCACAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500

Qy   7501 CGGTGTCATGTGGGACTATGATGGGAAGCAGCTCACACGCTCACAGGTGCCTTCTGCTAAGTC 7560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7501 CGGTGTCATGTGGGACTATGATGGGAAGCAGCTCACACGCTCACAGGTGCCTTCTGCTAAGTC 7560

Qy   7561 CCACATCACTGGCCTTCGGGCACTGATGTTCGTTCTGAGCAGCCCCGCAAGGCTGTTCT 7620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7561 CCACATCACTGGCCTTCGGGCACTGATGTTCGTTCTGAGCAGCCCCGCAAGGCTGTTCT 7620

Qy   7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680

Qy   7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACAACAACCAACAAAGAAAACCCCCAAG 7740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACAACAACCAACAAAGAAAACCCCCAAG 7740

Qy   7741 GCTTATCTCGTACCCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGT 7800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7741 GCTTATCTCGTACCCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGT 7800

Qy   7801 TGCTCCTGACGTAGTTAAAGCTTGTCATGGGAGATGCGTACGGGTTTGTAGATCCACGTAC 7860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7801 TGCTCCTGACGTAGTTAAAGCTTGTCATGGGAGATGCGTACGGGTTTGTAGATCCACGTAC 7860

Qy   7861 CCGTGTCAAGAGGACGAGGTCTTCTGTGATGTCGGTCACCCGATCGGCCACATGCGATAC 7920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7861 CCGTGTCAAGAGCCGAGGTCTTCTGTGATGTCGGTCACCCGATCGGCCACATGCGATAC 7920

Qy   7921 AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTCGGAGACAGACATCTACTC 7980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7921 AGTGTGTTTTGACAGTACCATCACACCCGAGGATATCATGGTCGGAGACAGACATCTACTC 7980

Qy   7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCATTCGCACCATTGCCGAGGCAGTTATA 8040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCATTCGCACCATTGCCGAGGCAGTTATA 8040

Qy   8041 CCGTGGAGGACCGAAGGTTATGATGGCCTGGAGGATCGGCGGATATCCGAGGTGTTAGGTC 8100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8041 CCGTGGAGGACCGAAGGTTATGATGGCCTGGAGGATCGGCGGATATCCGAGGTGTTAGGTC 8100

Qy   8101 TTCCGGGCGTCTATACTACCTCAAGTTGACCTGGCTGGAAGTGTAAATGC 8160
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8101 TTCCGGGCGTCTATACTACCTCAAGTTGACCTGGCTGGAAGTGTAAATGC 8160
```

```
Qy   8101 TTCCGGCGGTCTATACTACCTCAAGTTCAACAGTTGACCTGGCTGGAAGGTAAATGC 8160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8101 TTCCGGCGGTCTATACTACCTCAAGTTCAACAGTTGACCTGGCTGGAAGGTAAATGC 8160

Qy   8161 TGCAGCCGAACAGGCTGGCAGCGGCTGGAAGAAGAACCCGCTTCCTTATTTGCGGCGAT GATTGCAC 8220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8161 TGCAGCCGAACAGGCTGGCAGCGGCTGGAAGAAGAACCCGCTTCCTTATTTGCGGCGAT GATTGCAC 8220

Qy   8221 CGTAATTTGGAAGAGCGCCGGAGCAGATGCAGACAAACAAGCAATGGCGTGTCTTTGCTAG 8280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8221 CGTAATTTGGAAGAGCGCCGGAGCAGATGCAGACAAACAAGCAATGGCGTGTCTTTGCTAG 8280

Qy   8281 CTGGATGAAGGTGATGGGTGCACCACAGATTGTGTGCCTCAACCCAAATACACAGTTTGGA 8340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8281 CTGGATGAAGGTGATGGGTGCACCACAGATTGTGTGCCTCAACCCAAATACACAGTTTGGA 8340

Qy   8341 AGAATTAACATCATGCTCATCAAATGTCGTATCCCCCTTGGCAGGTGTCTCGCCGAGGGTCTGGG 8400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8341 AGAATTAACATCATGCTCATCAAATGTCGTATCCCCCTTGGCAGGTGTCTCGCCGAGGGTCTGGG 8400

Qy   8401 CTACTTTCTTACAAGAGATCCTGCGTCTGCGGGATTGGGTATCTAATACATCAATCACTCTCT 8460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8401 CTACTTTCTTACAAGAGATCCTGCGTCTGCGGGATTGGGTATCTAATACATCAATCACTCTCT 8460

Qy   8461 ATACAACCCCAGTTGCTGCGTCTGCGGGTATTGGAGGCTTTCTCGGTGCGGTGTCTACCAACGCTGA 8520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8461 ATACAACCCCAGTTGCTGCGTCTGCGGGTATTGGAGGCTTTCTCGGTGCGGTGTCTACCAACGCTGA 8520

Qy   8521 TAGCCGTGTGTTGGCTGTGTCCATTTCATGGAGCAGATGTCTCTTTGAGGACAAACTTCCCGA 8580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8521 TAGCCGTGTGTTGGCTGTGTCCATTTCATGGAGCAGATGTCTCTTTGAGGACAAACTTCCCGA 8580

Qy   8581 GACTGTGACCTTTGACTGTGTATGGGAAAAATTATACGGTGCCTGTGAGAAGATCTGCCCCAG 8640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8581 GACTGTGACCTTTGACTGTGTATGGGAAAAATTATACGGTGCCTGTGAGAAGATCTGCCCCAG 8640

Qy   8641 CATCATTGCTGGTGCGTGCACGGTATTGAGGCTTTCTCGGTGCGGTACACCAACGCTGA 8700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8641 CATCATTGCTGGTGCGTGCACGGTATTGAGGCTTTCTCGGTGCGGTACACCAACGCTGA 8700

Qy   8701 GATCCTCAGATTCCCAATCACTAACAAGACATGACCATGCCCCCCTGCGAGCCTGGCG 8760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8701 GATCCTCAGAGTTCCCAATCACTAACAAGACATGACCATGCCCCCCTGCGAGCCTGGCG 8760

Qy   8761 AAAGAAAGCCAGGCGGGTCCTGCCAGCGCCAAGAGGCGTGGCGGAGCACACGCAAATT 8820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8761 AAAGAAAGCCAGGCGGGTCCTGCCAGCGCCAAGAGGCGTGGCGGAGCACACGCAAATT 8820

Qy   8821 GGCTTCGCGCTTCCTTCTGCGCATGCTACATCTAGACCTCACAGATTTGGATAAGACGAG 8880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8821 GGCTTCGCGCTTCCTTCTGCGCATGCTACATCTAGACCTCACAGATTTGGATAAGACGAG 8880

Qy   8881 CGTGGCTCGGTACACCACTTTCAATTATTGGATGTGATTCTCCCCCGGAGGGGGATGTGTT 8940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8881 CGTGGCTCGGTACACCACTTTCAATTATTGGATGTGATTCTCCCCCGGAGGGGGATGTGTT 8940

Qy   8941 TATTACACCAGAAGAGATTGCAGAAGTTCCTTGTGAAGTTTGGCTGTCATTGTTTT 9000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8941 TGTTACACCAGAAGAGATTGCAGAAGTTCGTGAAGTATTTGGCTGTCATTGTTTT 9000

Qy   9001 TGCCCTAGGGCTCATTGCGTGTTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9001 TGCCCTCAGGGCTCATTGCGTGTTGGACTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060

Qy   9061 CTAACAG---------TTTTTTTTTTTTTTTTT 9116
          ||||||                           
Db   9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTT 9120

Qy   9117 GGGCTTAACGACCCCGC 9133
          |||||||||||||||||
Db   9121 GGGCTTAACGACCCCGC 9137
```

RESULT 15

```
AR494275
LOCUS       AR494275           9143 bp    DNA      linear   PAT 15-MAY-2004
DEFINITION  Sequence 393 from patent US 6720166.
ACCESSION   AR494275
VERSION     AR494275.1  GI:47267264
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 9143)
AUTHORS     Simons,J.N., Pilot-Matias,T.J., Dawson,G.J., Schlauder,G.G.,
            Desai,S.M., Leary,T.P., Muehoff,A.S., Erker,J.C., Buijk,S.L. and
            Mushahwar,I.K.
TITLE       Non-a, non-b, non-c, non-c, non-d, non-e hepatitis reagents and
            methods for their use
JOURNAL     Patent: US 6720166-A 393 13-APR-2004;
FEATURES             Location/Qualifiers
source          1..9143
                /organism="unknown"
                /mol_type="genomic DNA"
ORIGIN

Query Match      96.4%;   Score 9059.8;   DB 6;   Length 9143;
Best Local Similarity   99.6%;   Pred. No. 0;
Matches 9096;   Conservative    0;   Mismatches   37;   Indels    4;   Gaps    1;
```

```
Qy         1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGAGCACCCCCCTAG   60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGAGCACCCCCCTAG   60

Qy        61 CAGGGGCGTGGGGGATTCCCCTGCCGTCTGCAGAAGGGTGGAGCCAACCACCTTAGTAT  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 CAGGGGCGTGGGGGATTCCCCTGCCGTCTGCAGAAGGGTGGAGCCAACCACCTTAGTAT  120

Qy       121 GTAGGCGGCGGCGGGGACTCATGAGACGCTCGCGTGATGATGACAAGCGCCAAGCTTGACTTGACTGGATGGC  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GTAGGCGGCGGCGGGGACTCATGAGACGCTCGCGTGATGATGACAAGCGCCAAGCTTGACTTGACTGGATGGC  180

Qy       181 CCTGATGGGCGTTCATGGGTTCGGTGGTGGTGGCGCTTTAGGCAGCCTCCACGCGCCACCA  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CCTGATGGGCGTTCATGGGTTCGGTGGTGGTGGCGCTTTAGGCAGCCTCCACGCGCCACCA  240

Qy       241 CCTCCCAGATAGAGCGGCGGCGACTGTAGGGAAGAACCGGGACCGTCACTACCAAGGACG  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CCTCCCAGATAGAGCGGCGGCGACTGTAGGGAAGAACCGGGACCGTCACTACCAAGGACG  300

Qy       301 CAGACCCTCTTTTGAGTCGTAGACCGGCAAGAACGCGCAAGACGCAGAGCGCAAGCTTCATATCCTGTGTCCATTAAAAC  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 CAGACCCTCTTTTGAGTCGTAGACCGGCAAGAACGCGCAAGACGCAGAGCGCAAGCTTCATATCCTGTGTCCATTAAAAC  360

Qy       361 TGGGATGGTTGGGGTTAGCCATCATCACCGCCTCCGAGCCTTGGCCTGTTTGCCAGC  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 TGGGATGGTTGGGGTTAGCCATCATCACCGCCTCCGAGCCTTGGCCTGTTTGCCAGC  420

Qy       421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACC  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACC  480

Qy       481 TGCGCCCAGAACGCGCAAGAACAACGAGCAAAGCGCAAAGTCCAGGCGCTTGGCTGCCAGC  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 TGCGCCCAGAACGCGCAAGAACAACGAGCAAAGCGCAAAGTCCAGGCGCTTGGCTGCCAGC  540

Qy       541 ATCTGTTGAAAGGGACAACGACCAAAGTCCGGCGGATGCTCGGCCTTCGCAA  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 ATCTGTTGAAAGGGACAACGACCAAAGTCCGGCGGATGCTCGGCCTTCGCAA  600

Qy       601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCCAGC  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCCAGC  660

Qy       661 TCATGGTTGGGACGGCCAAGACCCTGCCCATAAGTCTCGCAATCTTGGAATCTTCTGGA  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 TCATGGTTGGGACGGCCAAGACCCTGCCCATAAGTCTCGCAATCTTGGAATCTTCTGGA  720
```

```
Qy       721 TTACCCTTTGGGGTGGATTGGTGATGTGATGTTACAACTCACACACCTTAGTAGGCCGCTCGT  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 TTACCCTTTGGGGTGGATTGGTGATGTGATGTTACAACTCACACACCTTAGTAGGCCGCTCGT  780

Qy       781 GGCAGGAGCGGTCGTTCGACCAGTCTGCCAGATAGTACGCTTGCTGGAGGATGGAGTCAA  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 GGCAGGAGCGGTCGTTCGACCAGTCTGCCAGATAGTACGCTTGCTGGAGGATGGAGTCAA  840

Qy       841 CTGGGCTACTGGTTGGTTCGGTGCTCACCTTTTGTGGTATGTCTGCTATCTTGGGCCTG  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 CTGGGCTACTGGTTGGTTCGGTGCTCCACCTTTTGTGTGATGTCTGCTATCTTTGGGCCTG  900

Qy       901 TCCCTGTAGTGGGGCGGGGTCACTGACCAGACAAATACCACCATTCCTGACCAATTG  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 TCCCTGTAGTGGGGCGGGGTCACTGACCAGACAAATACCACCAATCCTGACCAATTG  960

Qy       961 CTGCCAGGTAATCAGTTATCATTGTCTCCTTCCACTTGCCTACACGAGCCTGGTTG  1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 CTGCCAGCGTAATCAGTTATCATTGTCTCCTTCCACTTGCCTACACGAGCCTGGTTG  1020

Qy      1021 TGTGATCTGTGCGGACGAGTGCTGCGGTTCCGCCAATCCGTACAATCTTCACACCTTCCAA  1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 TGTGATCTGCGGGACAGAGTGCTGCGGTTCCGCCAATCCGTACAATCTTCACACCTTCCAA  1080

Qy      1081 TTGGACTGGCACGGACTCCTTCTTGGCTGACCACATTGATTTGTTATGGGGCGCTCTTGT  1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 TTGGACTGGCACGGACTCCTTCTTGGCTGACCACATTGATTTGTTATGGGGCGCTCTTGT  1140

Qy      1141 GACCTGTGAGCGCCTTGACATTGGTGAGTTGTGTGGTGTGTATTAGTCGGTGACTG  1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 GACCTGTGACGGCCCTTGACATTGGTGAGTTGTGTGGTGTGTATTAGTCGGTGACTG  1200

Qy      1201 GCTTGTCAGGCACTGGCTTAATCACATAGAACTTCAATGAAACTGGTACTTGTTACCTGA  1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 GCTTGTCAGGCACTGGCTTAATCACATAGAACTTCAATGAAACTGGTACTTGTTACCTGA  1260

Qy      1261 AGTGCCACTGGAATAGATCCTGGGTTCCTAGGGTTCACAAGGTTTATCGGGATGGCCGGCAAGGT  1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 AGTGCCACTGGAATAGATCCTGGGTTCCTAGGGTTCACAAGGTTTATCGGGATGGCCGGCAAGGT  1320

Qy      1321 CGAGGCTGTCATCTTCTTGACCAAAACTGGCTTCACAAGTACCATCACTCATTGCGACTAT  1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 CGAGGCTGTCATCTTCTTGACCAAAACTGGCTTCACAAGTACCATCACTCATTGCGACTAT  1380

Qy      1381 GTTTAGCAGTGCTCTCCTAGCGGTTGGCGCTTACACTAGAAGGGACCTTGGAAACCCCAT  1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 GTTTAGCAGTGCTCTCCTAGCGGTTGGCGCTTACACTAGAAGGGACCTTGGAAACCCTAT  1440

Qy      1441 GTGGTATCAGTTGCTCTCCTAGCGCTTATGCTTTACATAGAGTTTGCTCGCCCTTGATGATACCATGTCC  1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 GTGGTATCAGTTGCTCTCCTAGCGCTTATGCTTTACATAGAGTTTGCTCGCCCTTGATGATACCATGTCC  1500

Qy      1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGTCGCGCTTTTGATGATCCTGTACC  1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGTCGCGCTTTTGATGATCCTGTACC  1560

Qy      1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCGAAGTCATTGTTTTACAGTCCAAAGTGGAC  1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 TTTGCCACTCTTATTTGAGTGAGAATGTGTCGAAGTCATTGTTTACAGTCCAAAGTGGAC  1620

Qy      1621 CAGGCCTATCACTCCTAGAGTATAACAACTCTTGGTACCCCCATACACAATCCTGG  1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 CAGGCCTGTCACTCCTAGAGTATAACAACTCCTATACCCCTATACACAATCCTGG  1680

Qy      1681 TGGCGAGGGGATGTATGGTAAATTCAAAAATAACACATGGGGTTGCTCCGTATTCGCAA  1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 TGGCGAGGGGATGTATGGTAAATTCAAAAATAACACATGGGGTTGCTCCGTATTCGCAA  1740

Qy      1741 TGTGCCATCGTACTGCACTATGGGCCACTATGGGCACTGGAACGACACTCGCAACACTTA  1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1741 TGTGCCATCGTACTGCACTATGGGCCACTATGGGCACTGGAACGACACTCGCAACACTTA  1800
```

Qy 1801 CGAAGCATGCGGTGTAACACCGATGGCTAACAACCGCATGGCACAACGGCTCAGCCCTGAA 1860
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 CGAAGCATGCGGTGTAACACCGATGGCTAACAACCGCATGGCACAACGGCTCAGCCCTGAA 1860

Qy 1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC 1920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC 1920

Qy 1921 AGGCCATTTGTATTTGAGGGATCAGATACCCTATAGTTACTTTTATGACCCTGTGAA 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921 AGGCCATTTGTATTTGAGGGATCAGATACCCTATAGTTACTTTTATGACCCTGTGAA 1980

Qy 1981 TTCCACTTCCTACCACCGGAGAGGTGGGCGTAGGTGCCCGGTACCCACCTGTGTACG 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981 TTCCACTTCCTACCACCGGAGAGGTGGGCGTAGGTGCCCGGTACCCACCTGTGTACG 2040

Qy 2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTACCACAGG 2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTACCACAGG 2100

Qy 2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC 2160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC 2160

Qy 2161 TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGTGTGCTAATTCTGTTGGGTTGTGTGG 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161 TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGTGTGCTAATTCTGTTGGGTGTGTGG 2220

Qy 2221 CAGCAAGTATCTTATTTTAGCCTCTGTGCTTCCCTTTGTTTTGTTTGGGCGCGCTTC 2280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221 CAGCAAGTATCTTATTTTAGCCTCTGTGCTTCCCTTTGTTTTGTTTGGGCGCGCTTC 2280

Qy 2281 TGGTTACCCTTTGCCTGTCCTGCTCCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGT 2340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2281 TGGTTACCCTTTGCCTGTCCTGCTCCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGT 2340

Qy 2341 TTTGTCTAAAGCTCAAGTAGCTCCTTTTGCTTTGATTTTCTTCATCGTTGCTATCTCCG 2400
         |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db 2341 TTTGTCTAAAGCTCAAGTAGCTCCTTTTGCTTTGATTTTCTTCATCGTTGCTATCTCCG 2400

Qy 2401 CTGCAGGCTACGTTATGCTGCCCTTTAGGGTTTGTGCCCATGGCCTTGCGCGGCTTC 2460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401 CTGCAGGCTACGTTATGCTGCCCTTTAGGGTTTGTGCCCATGGCCTTGCGCGGCTTC 2460

Qy 2461 AACTTCTTTGTTGTGCAGCAGTCTGCCCAACCAGATTATGACTGGTGGTGCGGACTGCT 2520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2461 AACTTCTTTGTTGTGCAGCAGTCTGCCCAACCAGATTATGACTGGTGGTGCGGACTGCT 2520

Qy 2521 AGTGGCGAGGGTAGTTTGTGGGCCGGCCGGCCGTGAACCGCTAACCGTGGCTCACCGCATAGCTCTGCTTGT 2580
         ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db 2521 AGTGGCAGGGTAGTTTGTGGGCCGGCCGGCCGTGACCGCTGGTCCACCGTAGCTCTGCTTCAT 2580

Qy 2581 AGTTCCTTGGCCTCGGCTTTTAACCCTCTTGCTCATTGGTTACGCCTTCAGC 2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2581 AGTTCCTTGGCCTCGGCTTTTAACCCTCTTGCTCATTGGTTAACTCCCTATCT 2640

Qy 2641 TTTTGATACCGGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT 2700
         ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2641 TTTTGACACCGGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT 2700

Qy 2701 GTCTCGTTTGGCTTCTTTGCTCACTTGTTACCTCGCTGTGCTTTTAGTTAACTCCTATCT 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2701 GTCTCGTTTGGCTTCTTTGCTCACTTGTTACCTCGCTGTGCTTTTAGTTAACTCCTATCT 2760

Qy 2761 TTGGCAACGTTGGGAGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTTCCT 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761 TTGGCAACGTTGGGAGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGTTTCTCCT 2820

Qy 2821 TGTGCTGGTTGTTGTTCCCCGGTTGCGCACATATGACCGGCGTGGTGACTTTCTCGTGTGTCA 2880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2821 TGTGCTGGTTTGTTGTTCCCCGGTTGCGACATATGACACGCGGTGGTGACTTTTCGTGTGTCA 2880

Qy 2881 CGTAGCTCTTCTATGTTTAACATCCAGTGCAGCCATCCAGTCGTTCTTTGGGACTGACTCTAGGGT 2940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2881 CGTAGCTCTTCTATGTTTAACATCCAGTGTGCGCTCTTCGTTCTTTGGGACTGACTCTAGGGT 2940

Qy 2941 TAGGGCCCATAGAAATGTTGGTGCGTCTCGGAAAGTGTCATGCTTGGTATTCTCATTATGT 3000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2941 TAGGGCCCCATAGAATGTTGGTGCGTCTCGGAAAGTGTCATGCTTGGTATTCTCATTATGT 3000

Qy 3001 TCTTAAGTTTTCCTCTTAGTGTTGGTGAGAATGGTGTTTTCTATAAGCACTTGCA 3060
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3001 TCTTAAGTTTTCCTCTTAGTGTTGGTGAGAATGGTGTTTTCTATAAGCACTTGCA 3060

Qy 3061 TGGTGATGGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTGCAAGAGGCCATTTTTCCC 3120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3061 TGGTGATGGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTGCAAGAGGCCATTTTCCC 3120

Qy 3121 TTTTGAAGGCAAGGCAAGGGTCATAGGAGTCATAGGAGTGCCTTGGCGTGTGGGGACAC 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3121 TTTTGAAGGCAAGGCAAGGGTCATAGGAGTCATAGGAGTGCCTTGGCGTGTGGGGACAC 3180

Qy 3181 GGTTGATGGTTTGCCGCGTTCTCGGCGACCTTGTTTTCGCAGGGTTGGCTAT 3240
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3181 GGTTGATGGTTTGCCGCGTTCTCGGCGACCTTGTTTTCGCAGGGTTAGCTAT 3240

Qy 3241 GCCGCCAGATGGGGTGGGCCATTACCGCACCTTTTACGCTGCAGTGTCTCTCTGAACGTGG 3300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3241 GCCGCCAGATGGGGTGGGCCATTACCGCACCTTTTACGCTGCAGTGTCTCTCTGAACGTGG 3300

Qy 3301 CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC 3360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3301 CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC 3360

Qy 3361 TATCTTCAGATTAGGATCTCTGGCCACTAGCTACATGGATTTGTTGTTGACAACGTGTT 3420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3361 TATCTTCAGATTAGGATCTCTGGCCACTAGCTACATGGATTTGTTGTTGACAACGTGTT 3420

Qy 3421 GTATACTGCTCACCATGGCAGCAAGGGGCCCGGTTGGCTCATCCCACCACCATGTGGAGCTGG 3480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3421 GTATACTGCTCACCATGGCAGCAAGGGGCCCGGTTGGCTCATCCCACCACCATGTGGAGCTGG 3480

Qy 3481 CCCAATAACCGTTGACGCGCTAATGACCGGCGCCGCTAATGACATCTTATCAACCACCATGTGGAGCTCG 3540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3481 CCCAATAACCGTTGACGCGCTAATGACCGGCGCCCCCGATTCTGTGCTCTCGTGTGGGGTC 3540

Qy 3541 GTCCCTTACTCGGTGCTCTTGCGGGGAGACCAAGGGGTATCTGGTAACACGACTGGGGTC 3600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3541 GTCCCTTACTCGGTGCTCTTGCGGGGAGACCAAGGGGTATCTGGTAACGACTGGGGTC 3600

Qy 3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGGTGCGGGGCCCTTCCCAT 3660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGGTGCGGGGCCCTTCCCAT 3660

Qy 3661 GGCTGTTGCCAAGGGTTCTTCAGGTGCCATGGCCTTGTACGGCCTTCAGC 3720
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3661 GGCTGTTGCCAAGGGTTCTTCAGGTGCCATGGCCTTGTACGGCATGTTATTGG 3720

Qy 3721 GATGTTCACCGGCTGCTAGCAAATTCTGGCGGTTCAGTCAGTTAGGGTTAGGCCGTT 3780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3721 GATGTTCACCGGCTGCTAGCAAATTCTGGCGGTTCAGTCAGTTAGGGTTAGGCCGTT 3780

Qy 3781 GGTGTGTGCTGGATACCATCCCAGTACACAGCCACTCTTGATACAAAACCTAC 3840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3781 GGTGTGTGCTGGATACCATCCCAGTACACAGCCACTCTTGATACAAAACCTAC 3840

Qy 3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC 3900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC 3900

Qy 3901 CAAATACCACCTTTCTTACATGCAGGAGAGTATGAGGTCTTGGTCCTAAATCCCAGTGT 3960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3901 CAAATTACCACCTTTCTTACATGCAGGAGAGTATGAGGTCTTGGTCCTAAATCCCAGTGT 3960

Qy 3961 GGCTACAACGACATCAATGCCAAAGTACAATGCCAAAGTACACGCGACGTACGCGGCTCTAGGGT 4020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db   3961 GGCTACAACAGCATCAATGCCAAAGTACATGCACGCGACGTACGGCGTGAATCCAAATTG 4020

Qy   4021 CTATTTAATGCAAATGTACCAACACAGGGCGCTTCACTTACGTTACGTACAGCACATGGCAT 4080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 CTATTTTAATGCAAATGTACCAACACAGGGCGCTTCACTTACGTTACGTACAGCACATGGCAT 4080

Qy   4081 GTACCTGACCGGAGCATGTTCCCGGAACTATGATGTAATCATTGTGACGAATGCCATGC 4140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 GTACCTGACCGGAGCATGTTCCCGGAACTATGATGTCATCATTGTGACGAATGCCATGC 4140

Qy   4141 TACCGATGCAACCACCGTGTTGGGCATTGGGCATTGGGAAAGTTCCTAACCGAAGCTCCATCCAAAAA 4200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 TACCGATGCAACCACCGTGTTGGGCATTGGGCATTGGGAAAGTTCTAACCGAAGCTCCATCCAAAAA 4200

Qy   4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCTGGAGTAATCCTACCACCACATGC 4260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCTGGAGTAATCCTACCACACCACATGC 4260

Qy   4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCTTTCATGGAAAAAGAT 4320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCTTTCATGGAAAAAGAT 4320

Qy   4321 TAAGGAGGAAAATCTGAAGAAAGGGAGACACCTTATCTTTGAGGCTACCAAAAAAACACTG 4380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321 TAAGGAGGAAAATCTGAAGAAAGGGAGACACCTTATCTTTGAGGCTACCAAAAAAACACTG 4380

Qy   4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTACTATAGGGG 4440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTACTATAGGGG 4440

Qy   4441 ATGTGACATCTCAAAAATCCTGAGGGGCGACTGTGTAGTAGTTGCCACTGGATGCCTTGTG 4500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 ATGTGACATCTCAAAAATCCCGAGGGGCGACTGTGTAGTAGTTGCCACTGGATGCCTTGTG 4500

Qy   4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTGCAGCCTCATGGTAGAAGGCAC 4560
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTGCAGCCTCATGGTAGAAGGCAC 4560

Qy   4561 ATGCCATGGTTGACCTTGACCCTACTTTCACCATGGCCATGGGGTCGCGGGGTTTCAGC 4620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 ATGCCATGGTTGACCTTGACCCTACTTTCACCATGGCCATGGGGTCGCGGGGTCTCAGC 4620

Qy   4621 AATAGTTAAAGGCCAGCCGTAGGGGCCGTGGGAGAGCTGGCATATACTACTA 4680
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4621 AATAGTTAAAGGCCAGCCGTAGGGGCCGTGGGAGAGCTGGCATATACTACTA 4680

Qy   4681 CTATCGCACCCGGGGTTGTACCCCTTCGGGGTATGTGTTCGGTGTGCGCGGGGTTCGAGCCTT 4740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4681 CTATCGCACCCCAACCTGGGGTTGTACCCCTTCGGGGTATGTGTTCGGTGTGCGCGGGGTCTCC 4740

Qy   4741 CGACGCAGCCAAGGCACATGGTATGGTTGTCATCACACTACAACATGTGTCATCAGTATGGCTATGCTGCTCC 4800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 CGACGCAGCCAAGGCACATGGTATGGTTGTCATCACACTACAACATGTGTCATCAGTATGGCTATGCTGCTCC 4800

Qy   4801 CTATCGCACCCAACCTGGGGTTACCTGGGTTACCTGCGATAGAGGAGCAAATTTGGAC 4860
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 CTATCGCACCCAACCTGGGGTTACCTGGGTTACCTGCGATAGAGGAGCAAATTTGGAC 4860

Qy   4861 CTTTTCTATGGTCAACCCCGAACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA 4920
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861 CTTTTCTATGGTCAACCCCGAACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA 4920

Qy   4921 TTATGTTTGTGTCAGCGAGGCATGGTATGGCTCATCAGTATGGCTATGCTGCTCC 4980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 TTATGTTTGTGTCAGCGAGGCATGGTATGGCTCATCAGTATGGCTATGCTGCTCC 4980

Qy   4981 CAATGACGCACCACGGTGCACCCGGCTTGGCAGGGAGCCCGGCTTGTGGGGTTCGTG 5040
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 CAATGACGCACCACGGTGCACCCGGCTTGGCAGGGAGCCCGGCTTGTGGGGTTCGTG 5040

Qy   5041 GCGCTTGGACGGCGCTGTCCTGGCCCAGAGCCCAGCAGCGCCCAGCGCCCAGAGCCCAGCGAGCGCCCAGAGCCCAGAGATACCA 5100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041 GCGCTTGGACGGCGCTGTCCTGGCCCAGAGCCCAGCAGCGCCCAGCGCCCAGAGCCCAGCGAGCGCCCAGAGCCCAGAGATACCA 5100

Qy   5101 AATGTGCTTCACTGAAGTCAATACTTCTGGGACAGCCGCACTCGCTGTTGGCGTTGGAGT 5160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101 AATGTGCTTCACTGAAGTCAATACTTCTGGGACAGCCGCACTCGCTGTTGGCGTTGGAGT 5160

Qy   5161 GGCTATGGCTTATCTAGCATTGACACTTTGCGCGCCACTTGTGTGCGCGGTTGCTGGTC 5220
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161 GGCTATGGCTTATCTAGCATTGACACTTTGCGCGCCACTTGTGTGCGCGGTTGCTGGTC 5220

Qy   5221 TATTACATCAGTCCTACCGGTGCTACTGTCCGCCCCAGTGGTGACGAAGAAGAAATCGT 5280
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221 TATTACATCAGTCCTACCGGTGCTACTGTCCGCCCCAGTGGTGACGAAGAAGAAATCGT 5280

Qy   5281 GGAGGAGTGTGCATCATTCATTCCCTGGAGGCCATGGTTGCTGCAATTGACAAGCTGAA 5340
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5281 GGAGGAGTGTGCATCATTCATTCCCTGGAGGCCATGGTTGCTGCAATTGATAAGCTGAA 5340

Qy   5341 GAGTACAATCACCACAACTAGTCCTTCACATTGGAAACCGCCCTTGAAAAACTTAACAC 5400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5341 GAGTACAATAACCACAACTAGTCCTTCACATTGGAAACCGCCCTTGAAAAAACTTAACAC 5400

Qy   5401 CTTTCTTGGGCCTCATGCGAGCTACAATCCTTGCTCACAATCCTTGCTCGTGTGTTTAGT 5460
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5401 CTTTCTTGGGCCTCATGCGAGCTACAATCCTTGCTCATCAGAGATTGCTGTGTGTTACTAC 5460

Qy   5461 CACTTTACCTGACAATCCCTTGCATCATGCGTGTTTGCTTTCATTGGGGGTATTACTAC 5520
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5461 CACTTTACCTGACAATCCCTTGCATCATGCGTGTTTGCTTTCATTGGGGGTATTACTAC 5520

Qy   5521 CCCACTACCTCACAAGATCAAAATGTTCCGTCGTCATTATTTGGAGGCGCAATTGCGTCCAA 5580
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5521 CCCACTACCTCACAAGATCAAAATGTTCCGTCGTCATTATTTGGAGGCGCAATTGCGTCCAA 5580

Qy   5581 GCTTACGACGGTAGAGGCGCACTGGCGTTCATGATGGCCGGGCGGCTGCGGGAACAGCTCT 5640
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5581 GCTTACGACGGTAGAGGCGCACTGGCGTTCATGATGGCCGGGCGGCTGCGGGAACAGCTCT 5640

Qy   5641 TGGTACATGGACATCGGTGGGTTTGTCTTGACAATGGCCGCCTAGGCGGCTACGCCGCCTC 5700
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5641 TGGTACATGGACATCGGTGGGTTTGTCTTGACAATGGCCGCCTAGGCGGCTACGCCGCCTC 5700

Qy   5701 ATCCACTGCTTGTTGTTGACATTTAAATGCTTGATTGTGAGTGCGCCACTATGGGATCAGCT 5760
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5701 ATCCACTGCTTGTTGTTGACATTTAAATGCTTGATTGTGAGTGCGCCACTATGGGATCAGC 5760

Qy   5761 TGCTGGTTAGTGCTTGACAACAGCAGGGCCCAGAGTGTGGGCGTCTGTGTCAGC 5820
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5761 TGCTGGTTAGTGCTTGACAACAGCAGGGCCCAGAGTGTGGGCGTCTGTGTCAGC 5820

Qy   5821 TGTTGCAATGTTGCTTGACAACAGGCCAGATCACTGGCCAACCAACAGACTTCTTAC 5880
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5821 TTGTGCAATGTTGCTTGACAACAGGCCAGATCACTGGCCAACCAACAGACTTCTTAC 5880

Qy   5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCCGTGACATCCG 5940
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCCGTGACATCCG 5940

Qy   5941 CAGGAAGATACTGGGCATTCGGAGGCATCTACCCCCCTGGAGTGTCATATCAGCTTGCAT 6000
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5941 CAGGAAGATACTGGGCATTCGGAGGCATCTACCCCCCTGGAGTGTCATATCAGCTTGCAT 6000

Qy   6001 CCGTTGGCTCCACACCCCGACGAGGATGATTGCGCGGCCTCATTGCTTGCGGTCTAGAGAT 6060
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6001 CCGTTGGCTCCACACCCCGACGAGGATGATTGCGCGGCCTCATTGCTTGCGGTCTAGAGAT 6060

Qy   6061 TTGGCAGTATGTGTGCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGTCTGGAGTTCA 6120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6061 TTGGCAGTATGTGTGCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGTCTGGAGTTCA 6120

Qy   6121 GAGCATGGTTAACATTCCTGTGTCCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC 6180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6121 GAGCATGGTTAACATTCCTGTGTCCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC 6180
```

Qy 7261 TTCGTGCAGCATGAGCTACACTTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TTCGTGCAGCATGAGCTACACTTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT

Qy 7321 TCTGTCTGCAACTCGGGCCCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGT 7380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TCTGTCTGCAACTCGGGCCCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGT

Qy 7381 GACTGAGCCGCGGGATGCGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GACTGAGCCGCGGGATGCGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCT

Qy 7441 GTTCCCCATCATACCAACAAGCAAGTGAGATTGGCTAAGGAGAAAAAGCTTCAAAAGTTGT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GTTCCCCATCATACCAACAAGCAAGTGAGATTGGCTAAGGAGAAAAAGCTTCAAAAGTTGT

Qy 7501 CGGTGTCATGTGGGACTATGATGAAGTAGCAGCTCACACGCCCTCCAGTCTCGCTAAGTC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CGGTGTCATGTGGGACTATGATGAAGTAGCAGCTCACACGCCCTCCAGTCTCGCTAAGTC

Qy 7561 CCACATCACTGGCCTTCGGGGCACTGATGTTCGTTCTGAGCAGCCCGCAAGGCTGTTCT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CCACATCACTGGCCTTCGGGGCACTGATGTTCGTTCTGAGCAGCCCGCAAGGCTGTTCT

Qy 7621 GGACTTGCAGAAGTGTGTCGAGGGAGGTGAAGACCCCCAGAAACCAACAACAAGAAAACCCCCAAG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GGACTTGCAGAAGTGTGTCGAGGGAGGTGAAGACCCCCAGAAACCAACAACAAGAAAACCCCCAAG

Qy 7681 AGTTCCAAAGGAGGAGAGTCTTCGTGAAGACGTGTTGAGAAGATGAGATGTGTTGAGAAGATGTACTGGTCAGGT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   AGTTCCAAAGGAGGAGAGTCTTCGTGAAGACGTGTTGAGAAGATGAGATGTGTTGAGAAGATGTACTGGTCAGGT

Qy 7741 GCTTATCTCGTACCCCCACCTTGAAATGAGATGTTGAGAAGATGTGTTGAGAAGATGTCAGT
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   GCTTATCTCGTACCCCCACCTTGAAATGAGATGTTGAGAAGATGTGTTGAGAAGATGTCAGT

Qy 7801 TGCTCCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTACGGGTTTGTGATCCACGTAC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TGCTCCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTACGGGTTTGTGATCCACGTAC

Qy 7861 CCGTGTCAAGCGTCTGTTGTCGATGTGGTCACCCGAGTCGGAGCCACATGCGCACATGCGATAC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CCGTGTCAAGCGTCTGTTGTCGATGTGGTCACCCGAGTCGGAGCCACATGCGCACATGCGATAC

Qy 7921 AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTGGAGACAGACATCTACTC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTGGAGACAGACATCTACTC

Qy 7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCATTCCACCATTGCGAGGGCAGTTATA
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCATTCCACCATTGCGAGGGCAGTTATA

Qy 8041 CGCTGGAGGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTAGGTC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CGCTGGAGGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTAGGTC

Qy 8101 TTCCGGTCTATACTACCTCAAGTTGACCTGCTGGCTGAAGGTAAATGC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TTCCGGTCTATACTACCTCAAGTTGACCTGCTGGCTGAAGGTAAATGC

Qy 8161 TGCAGCCGAACAGGCTGGCATGAAGAAGAACCCCGCTTCCGCTTCCGCGGGCATGATTGCAC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   TGCAGCCGAACAGGCTGGCATGAAGAAGAACCCCGCTTCCGCTTCCGCGGGCATGATTGCAC

Qy 8221 CGTAATTTGGAAGAGACCGAGCGCCGAGCAGATGCAGACAAACAAGCAGACAAACAAATACAGTTTGCTCGCTG
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CGTAATTTGGAAGAGACCGAGCGCCGAGCAGATGCAGACAAACAAGCAGACAAACAAATACAGTTTGCTCGCTG

Qy 8281 CTGGATGAAGGTGATGGGTCACCCACAAGATTGTGCCTCAACCCAAATACAGTTTGGA
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   CTGGATGAAGGTGATGGGTCACCCACAAGATTGTGCCTCAACCCAAATACAGTTTGGA

Qy 8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGAATTACCAAAAGTGCAAGCCCTTA
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 6181 CTGGGATTGGATCAGGTATGCTCCAAGCACGCTGTCTCCATGCGGTGCTGAACTCATCTTTC 6240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6181 CTGGGATTGGATCAGGTATGCTCCAAGCACGCTGTCTCCATGCGGTGCTGAACTCATCTTTC 6240

Qy 6241 TGTTGAGAAATGGTTTCAAAACTTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG 6300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6241 TGTTGAGAAATGGTTTCAAAACTTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG 6300

Qy 6301 AGGGGCTGTTCCAGTCAACGCTAGGCGTGTGGTGGCTAGACCGGACCCAACTGATTG 6360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6301 AGGGGCTGTTCCAGTCAACGCTAGGCGTGTGGTGGCTAGACCGGACCCAACTGATTG 6360

Qy 6361 GACTAGTTCGTCGTCAATTATGGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA 6420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6361 GACTAGTTCGTCGTCAATTATGGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA 6420

Qy 6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCCAAC 6480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCCAAC 6480

Qy 6481 CTTGAGAGCTGCAGTGGCCGTGACGGCGTACAGGTTCAGTGTTATCGGTGAGCGTGATACCCAA 6540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6481 CTTGAGAGCTGCAGTGGCCGTGACGGCGTACAGGTTCAGTGTTATCGGTGAGCGTGATACCCAA 6540

Qy 6541 AACTCCTTGACACATCTGCTTGCTGTGCACACCTGTGCGCATGCAACTTAATTGCGTGA 6600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6541 AACTCCTTGACACATCTGCTTGCTGTGCACACCTGTGCGCATGCAACTTAATTGCGTGA 6600

Qy 6601 GCTTCCCTTCGACGGTCACACACTCGGTGTGCGCATGCACACACTCGGTGTGCGCATGCGTGA 6660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6601 GCTTCCCTTCGACGGTCACACACTCGGTGTGCGCATGCACACACTCGGTGTGCGCATGCGTGA 6660

Qy 6661 TGCACTTGAGACAAATGACTGTAATTCCACAAACAACAACTCCTAGTGATGAGCCGCAGT 6720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6661 TGCACTTGAGACAAATGACTGTAATTCCATAAACAACAACTCCTAGTGATGAGCCGCAGT 6720

Qy 6721 GTCGGCTCTTGTTTTCAAAACGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC 6780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6721 GTCGGCTCTTGTTTTCAAAACGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC 6780

Qy 6781 AGCTGGGCGTTGACACCACCACCTGTTCTACAGTGCCAGCCCCCTACCTGCAGAGAGGTAGTGTAAGAAA 6840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6781 AGCTGGGCGTTGACACCACCACCTGTTCTACAGTGCCAGCCCCCTACCTGCAGAGAGGTAGTGTAAGAAA 6840

Qy 6841 GCGCCAGTTCCGGGCAAGAACTCGGTCGTCCGGCTTCGCTTACCTTGCCTCCCCCTCCGAGATCCGTCCC 6900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6841 GCGCCAGTTCCGGGCAAGAACTCGGTCGTCCGGCTTCGCTTACCTTGCCTCCCCCTCCGAGATCCGTCCC 6900

Qy 6901 AGGAGTGTCATGTCCTGCAACGAAGCCTGCAACGAGTGCAAACAGGCGGAGGCCCTTCAAACCT 6960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6901 AGGAGTGTCATGTCCTGCAACGAAGCCTGCAACGAGTGCAAACAGGCGGAGGCCCTTCAAACCT 6960

Qy 6961 CCCTTCCTTCCACCACCCTGTTCTCACCACCCTGTTCTCCATGGCCCATGCCCCGTTGGGAGCCGGGGTGA 7020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6961 CCCTTCCTTCCACCACCCTGTTCTCACCACCCTGTTCTCCATGGCCCATGCCCCGTTGGGAGCCGGGGTGA 7020

Qy 7021 GTGTAACCCTTTCACTGGTGTCATGTCGAATTGGATGTGCAAATGACCGGAAACAGGCGGAGGCCCTGATGA 7080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7021 GTGTAACCCTTTCACTGGTGTCATGTCGAATTGGATGTGCAAATGACCGGAAACAGGCGGAGGCCCTGATGA 7080

Qy 7081 TTTACCCAGTTACCCTCCAGCTTGGGCACTGGCCAGCCCCCGTTGGCAGATCGGGGAAAGGA 7140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7081 TTTACCCAGTTACCCTCCAGCTTGGGCACTGGCCAGCCCCCGTTGGCAGATCGGGGAAAGGA 7140

Qy 7141 GGCTACAACCGCTTCCAGCTACGCCGCAAACGGCTACAACGCCCGCCAAACGGCTACAACCGGCTACAACCGGCTACAACCGGTTGGTCAAC 7200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7141 GACTACAACCGCTTCCAGCTACGCCGCAAACGGCTACGCCCGCCAAACGGCTACGCCCGCCAAACGGCTACAACCGGTTGGTCAAC 7200

Qy 7201 TTCCACTCAGTCAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7201 TTCCACTCAGTCAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7260

```
Db   8341  AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA  8400

Qy   8401  CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTGCCGAGGGTCTGGG  8460
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8401  CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTGCCGAGGGTCTGGG  8460

Qy   8461  ATACAACCCCAGTGCGTGCGTGGATTGGGTATCTAATACATCACTACCCATGTTTGTGGGT  8520
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8461  ATACAACCCCAGTGCGTGCGTGGATTGGGTATCTAATACATCACTACCCATGTTTGTGGGT  8520

Qy   8521  TAGCCGTGTGTTGGCTGTCCATTTCATGGAGCAGAATGCTCTTTGAGGACAAACTTCCCGA  8580
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8521  TAGCCGTGTGTTGGCTGTCCATTTCATGGAGCAGAATGCTCTTTGAGGACAAACTTCCCGA  8580

Qy   8581  GACTGTGACCTTTGACTGGTATGGGAAAAATATACGGTGCCTGTAGAAGATCTGCCCAG  8640
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8581  GACTGTGACCTTTGACTGGTATGGGAAAAATATACGGTGCCTGTAGAAGATCTGCCCAG  8640

Qy   8641  CATCATTGCTCGTGTGCACGGTATTGAGGCTTTCTCCGTGGTGCGCTACACCAACGCTGA  8700
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8641  CATCATTGCTCGTGTGCACGGTATTGAGGCTTTCTCCGTGGTGCGCTACACCAACGCTGA  8700

Qy   8701  GATCCTCAGAGTTCCCAATCACTAACAGACATGACCATGCCCCCCTGCGAGCCTGGCG  8760
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8701  GATCCTCAGAGTTCCCAATCACTAACAGACATGACCATGCCCCCCTGCGAGCCTGGCG  8760

Qy   8761  AAAGAAAGCCAGGGCGGTCCTCGCCAGCGCCAAGAGGCGTGCGGAGCACACGCAAAATT  8820
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8761  AAAGAAAGCCAGGGCGGTCCTCGCCAGCGCCAAGAGGCGTGCGGAGCACACGCAAAATT  8820

Qy   8821  GGCTCGCTTCCTTCCTCTCTGGCATGCTACATCTAGACCTCTACCAGATTTGGATAAGACGAG  8880
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8821  GGCTCGCTTCCTTCCTCTCTGGCATGCTACATCTAGACCTCTACCAGATTTGGATAAGACGAG  8880

Qy   8881  CGTGGCTCGGTACACCACTTTCAATTATTGTGATGTTACTCCCCGAGGGGATGTGTT  8940
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8881  CGTGGCTCGGTACACCACTTTCAATTATTGTGATGTTACTCCCCGAGGGGATGTGTT  8940

Qy   8941  TATTACCACAGAGAAGATTGCAGAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTTT  9000
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8941  TGTTACACCACAGAGAAGATTGCAGAAGTTCTTGTGAAGTATTTGGCTGTCATTGTTTT  9000

Qy   9001  TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA  9060
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9001  TGCCCTAGGGCTCATTGCTGTTGGACTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA  9060

Qy   9061  CTAACAG----TTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC  9116
                  ||||||  |||||||||||||||||||||||||||||||||||||||||
Db   9061  TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC  9120

Qy   9117  GGGCTTAACGACCCCGC  9133
                  ||||||||||||||||
Db   9121  GGGCTTAACGACCCCGC  9137
```

Search completed: October 30, 2005, 18:15:31
Job time : 25704 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2005, 09:45:50 ; Search time 17780 Seconds
(without. alignments)
20121.826 Million cell updates/sec

Title: US-10-009-002-1
Perfect score: 9399
Sequence: 1 accacaacactccagttg.........cccgcttggaattaaaaact 9399

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 34239544 seqs, 1903213470 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
	1	53.4	0.6	1101	9	CNS016HG	AL106750 Drosophil
	2	52.4	0.6	1101	9	CNS00LT2	AL078714 Drosophil
c	3	51.8	0.6	988	9	CNS008J9	AL052147 Drosophil
	4	49.2	0.5	939	9	CNS0150M	AL105712 Drosophil
	5	48.4	0.5	997	9	CNS005TE	AL060767 Drosophil
	6	48	0.5	922	9	CNS0073W	AL066784 Drosophil
c	7	47.4	0.5	937	9	CNS006ST	AL065880 Drosophil
	8	46.2	0.5	884	9	CNS006U0	AL065923 Drosophil
	9	46	0.5	895	9	CNS0071A	AL066286 Drosophil
c	10	46	0.5	947	9	CNS0047J	AL066360 Drosophil
	11	45.4	0.5	841	9	AG484307	AG484307 Mus muscu
	12	45	0.5	919	9	CNS0065S	AL065856 Drosophil
	13	45	0.5	1201	9	CNS0164Y	AL066100 Drosophil
c	14	43.8	0.5	925	9	CNS0091P	AL053013 Drosophil
	15	43.4	0.5	902	9	CNS006QP	AL065804 Drosophil
	16	43.2	0.5	910	9	CNS006QN	AL065629 Drosophil
c	17	43.2	0.5	1101	9	CNS017JU	AL108132 Drosophil
c	18	43	0.5	734	4	BG426743	BG426743 602493292
	19	43	0.5	1159	9	CNS015XR	AL106041 Drosophil
c	20	42.8	0.5	340	1	AI098746	AI098746 ue38a01.y
	21	42.8	0.5	847	7	CK128702	CK128702 AGENCOURT
c	22	42.4	0.5	474	4	BM522177	BM522177 ESSU0714
c	23	42.4	0.5	967	7	CR293420	CR293420 CR293420
c	24	42.4	0.5	979	9	CNS016BU	AL106548 Drosophil
c	25	42.2	0.4	540	9	CE184530	CE184530 tigr-gss-
c	26	42.2	0.4	570	5	BM929584	BM929584 UI-E-EJ1-
c	27	42.2	0.4	573	4	BM683407	BM683407 UI-E-EJ1-
c	28	42.2	0.4	776	7	CV059103	CV059103 BNEL44d5
	29	42	0.4	721	9	CC803018	CC803018 ih10c06.g
c	30	42	0.4	866	9	AG126308	AG126308 Pan trogl
c	31	42	0.4	885	4	BM416014	BM416014 OP21097 M
	32	42	0.4	1100	9	CNS014RP	AL104527 Drosophil
	33	42	0.4	1190	5	BX335937	BX335937 Drosophil
c	34	41.8	0.4	704	9	AG044304	AG044304 Pan trogl
c	35	41.8	0.4	1022	9	CNS04X27	AL311128 Tetraodon
	36	41.8	0.4	1072	9	CNS017PF	AL108333 Drosophil
	37	41.6	0.4	604	4	BJ483635	BJ483635 Drosophil
c	38	41.6	0.4	610	4	BJ466204	BJ466204 Drosophil
	39	41.6	0.4	909	9	CNS00JTL	AL076720 Drosophil
c	40	41.6	0.4	939	9	CNS00CNG	AL059400 Drosophil
c	41	41.6	0.4	991	9	BG967437	BG967437 602833555
	42	41.4	0.4	1101	9	CNS00Z1D	AL097099 Drosophil
c	43	41.2	0.4	1101	9	CNS012S8	AL101954 Drosophil
c	44	41.2	0.4	1109	9	CL059484	CL059484 CH216-90B
c	45	41.2	0.4	1200	3	CR653442	CR653442 Tetraodon

ALIGNMENTS

RESULT 1
CNS016HG
LOCUS   CNS016HG   1101 bp   DNA   linear   GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN16J15 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL106750
VERSION   AL106750.1 GI:5623622
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1101)
AUTHORS   Genoscope.
TITLE   Direct Submission
JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES   Location/Qualifiers
     source   1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN16J15"
                     /clone_lib="DrosBAC"
                     /plasmId="pBeloBAC11"
                     /note="end : SP6"
ORIGIN

Query Match   0.6%; Score 53.4; DB 9; Length 1101;
Best Local Similarity   18.4%; Pred. No. 0.0025;
Matches   73; Conservative 161; Mismatches 163; Indels   0; Gaps   0;

Qy   2227 GTATCTTATTTAGCCTACCTCGTTACTTGTCCCTTTGTTTGGGCGCCGCCTTCTGGTTA 2286
                ::|:|: |||| :|||:: ||| :|:|| :|:|:|::| |:|||:||:|||::||
Db   686 KYTYTBTCYKYGCTSYVYTCCCVKYTYTCYGTTKTSTTCSGGSCGSBCGGGTCCCCG 745

Qy 2287 CCCTTTGGCTCCTGCTCCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGTTTTGTC 2346
         : ||:: ||: :|| :|: : ||| :   :|| ||||| ||:: :: ::
Db   746 YTTTTTSYCTCBCSYGCSYCYYKTKGGCTCGCCSGBGGCGGCGKSTSTGBGGGSTKKKKSKSGY 805

Qy 2347 TAAAGCTTCAAGTAGCTCCTTTTGCTTTGATTTTCTTCATCTGTTGCTATCTCCGCTGCAG 2406
         :| |:|::||::|||::|||||||:::: ::
Db   806 GBGKTBTKTTTKTTTTRKTSTBSSBWTKTTSTTBSSBBTTATTTTTBSSTSSSTSCTSTBB 865

Qy 2407 GCTACGTTATGCTGCCCTTTAGGGTTGTGCCCATGCTGCCCCTAACCTT 2466
         :: :::: :|::: :|:|||::|:| ||||:::|::::|:
Db   866 BBCTSKSBSSTSTBSSBSSSTTTBSTSTSTSSBSYKYBTSSSSTSTSTSYSWCCCBCCYB 925

Qy 2467 CTTTGTTGCAGCAGCTGCTGCCCAACCAGATTATGACTGGTGGGTGCCGACTGCTAGTCGC 2526
         |:|::| ::|:: ::::| :: :|: ||:| |:|:
Db   926 CYSTBYTYSYSYCYCTBCBYSCTSTSCTTTBYKSTSBSYSTSTSYDYTTBSCBYSTBC 985

Qy 2527 AGGGTTAGTTTTTGTGGGCCGGCCGTAACCGGTCGTCACCGCATAGCTCTGCTTGTAGGTCC 2586
         :: :::: |: ||:::|:| ::: ::: :::
Db   986 TSGBASASTBTSSSBBSSSSSSCTSBBSSSSKSSSCSASTSTSTSTBCSSYYBCTSTSTSBS 1045

Qy 2587 TTGGCCCTCGGTAGCGCGTTTAACCCCTTGCATTG 2623
         :: :::: |:: ||:| :::|:
Db 1046 BCSTSYTYSBSSTSASBTTSYTBABTKSTBSBRBSWB 1082

Qy 2673 TACCACCTGTAGTAGCATTAGTTGTCATGTCCGTTTGCTTCATGTCTTTTGCTCACTTGTTAC 2732
         : |:::: ::| |: || :: ::: :: :::
Db   706 KAAAADKWDAKKWDGAKKATKTKKKKDKKAAWAAADKDRKKGKKGKGKKKGKKKK 765

Qy 2733 CTCGCTGTCTTTAGTTAACTCCTATCTTTGGCAACGTTGGGAGAATTCGTTTGGAACG 2792
         :: :::| :::|:: ::| |: :::::: ||:::|:
Db   766 KKKKKGKGGWGKKGKKKAGDDDAKDKTKKKKAATTTKKKKKGKKKKKKAAKKKKAADR 825

Qy 2793 TTACTAAGACCGGAGAGGTTTTTCCTTGTGCTGGTTGTGTTCCCGGTGCGACATATG 2852
         |:|: |||| ||:|::|| |:||::|::
Db   826 TKKTKWDAAAAAAAKKTTKDKGKKKKTKKKTTKKKKKGGGKKDAAAKKKKGTK 885

Qy 2853 ACGGCGTGGTGACTTTCTGTGGTGTCACGTAGCTCTTCTATGTTAACATCCAGTGCAG 2912
         |:|::| ::|::: :|:: ::
Db   886 KKKKKGKGGKGKGGGGGKGGKDDDAAAAKKKGKKKKKGKGGGGGKGKGGKK 945

Qy 2913 CATCGTTCTTTGGGACTGACTCTAGGGTTAGGGCCCATAGAAGTGTTGGTGCGTCTCGGAA 2972
         : ::::::: ||::|::||| ::|::| :|||
Db   946 KTKKTKKKKKKKKKADAAKGKTKKRAAADAAAADWTDAATKKKKKKKKDKKKKTKKKDA 1005

Qy 2973 AGTGTCATGCTTCGTGTATTCTCATTATGTCTTAAGTTTTTCCTCTTAGTGTTTGGTGAGA 3032
         :: :::|: ::| ||: ::| ::
Db 1006 AAAATKKKKKKTTKKKKKTTTADKGKKKKKKAADDKKKKK 1065

Qy 3033 ATGGTGTGTTTTTCTATAAGCACTTGCATGGTGATG 3068
         ::::: :|::: ::: :::
Db 1066 TKKKKKKKKKKTKKKTTTADKGKKKKKKAADDKKKKK 1101

RESULT 2
CNS00LT2
LOCUS   CNS00LT2   1101 bp   DNA   linear   GSS 14-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC:
            BACR48P19 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL078714
VERSION   AL078714.1 GI:5102004
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Oseegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
     source   1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone_lib="RPCI-98"
                     /clone="BACR48P19"
                     /note="end : TET3"
ORIGIN

Query Match   0.6%; Score 52.4; DB 9; Length 1101;
Best Local Similarity   12.6%; Pred. No. 0.005;
Matches   50; Conservative 182; Mismatches 164; Indels   0; Gaps   0;

RESULT 3
CNS008J9/c
LOCUS   CNS008J9   988 bp   DNA   linear   GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR17M13 of RPCI-98 library from Drosophila melanogaster (fruit
            fly) genomic survey sequence.
ACCESSION   AL052147
VERSION   AL052147.1 GI:4933797
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 988)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Oseegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
     source   1..988
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR17M13"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"

ORIGIN

Query Match   0.6%;   Score 51.8;   DB 9;   Length 988;
Best Local Similarity   47.0%;   Pred. No. 0.0073;
Matches 111;   Conservative 7;   Mismatches 118;   Indels 0;   Gaps 0;

Qy   2809 GAGGTTTTCCTTGTGCTGGTTGTTCCCCGGTGCGACATATGACGCGCTGGTGACTTT 2868
            ||||||||||||| ||| |||| |||||||||||||||||||||
Db   254 GTGCTTTTTCTTTCTTCCCCGTTTTCTTCCGGCGCCNCTGCGTNCGNCGTGNTTCTCNCTTT 195

Qy   2869 CTGTGTGTGTCACGTAGCTCTTCTATGTTAACATCCAGTGCAGCATCGTTCTTTGGGAC 2928
            |||||||| |||| || ||| : ||
Db   194 TGGTGTGTTGCCGTCGGTGGYTKGTTGGTYTTGTGTTTRCCCCGGGGGGGCCTGGTCGGGGGTKG 135

Qy   2929 TGACTCTAGGGTTAGGGCCCATAGAATGTGTGCGTCTCGGAAAGTGTCATGCTTGGTA 2988
            |||||| ||||||||| : |:|:
Db   134 TAGCTGNCGGGGGGGGGGTTCCTCNCTTTGGTGTGTCBCCGCYNNKGKYTGTTTTTGTG 75

Qy   2989 TTCTCATTATGTTCTTAAGTTTTCCTTCTTAGTGTTTGGTGAGAATGGTGTGTTT 3044
            :||||| |||| ||| ||| ||||
Db   74 GTTTGGTTTTTGTTGTGTSTTTTTAATCTTTGGGTTTTGTGTNNNNTGTNTTTTCT 19

RESULT 4
CNS015OM
LOCUS   CNS015OM   939 bp   DNA   linear   GSS 26-JUL-1999
DEFINITION   Drosophila melanogaster genome survey sequence SP6 end of BAC
              BACN14M12 of DrosBAC library from Drosophila melanogaster (fruit
              fly), genomic survey sequence.
ACCESSION   AL105712
VERSION   AL105712.1   GI:5617892
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
              Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
              Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
              Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 939)
AUTHORS   Genoscope.
TITLE   Direct Submission
JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
              collaboration with the European Drosophila Genome Project (EDGP) -
              http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
              library (Dros BAC) was made by Alain Billaud at CEPH (Centre
              d'Etude du Polymorphisme Humain) with funding provided by a MRC
              project grant. The DNA was prepared from embryos by Alain Bucheton
              and Genevieve Payan. It has been constructed in the vector
              pBeloBAC11.
FEATURES   Location/Qualifiers
     source   1..939
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN14M12"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
ORIGIN

Query Match   0.5%;   Score 49.2;   DB 9;   Length 939;
Best Local Similarity   40.5%;   Pred. No. 0.044;
Matches 83;   Conservative 34;   Mismatches 88;   Indels 0;   Gaps 0;

Qy   8886 CTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGAGGGGGATGTGTTTATTA 8945
            |||:||| |||||||||||||||||| |||| ||| || : :
Db   723 CTTGGWAMAAATATTTTTTTTTTTTTWWWTTTWTTTTSTCMGGMACCGAATACGTATMAYC 782

Qy   8946 CACCACAGAGAAGATTGCAGAAGTTCCTTGTGAAGTATTGGCTGTCATTGTTTTTGTCCC 9005
            ||| :||| :| ::||| | | ||| : || ||| ||| |||
Db   783 CCCTWACCMSAARCTTAGTTTTKAWAACCCTAARAAAATTDFDAAAWTTYYKTCGATAT 842

ORIGIN

Query Match   0.6%;   Score 51.8;   DB 9;   Length 988;

Qy   9006 TAGGGCTCCATTGCTGTGGATTAGCCATCAGTGAACCCCCAAATTCAAAATTAACTAC 9065
            |:||||:|:|| |||:|: | ::|| ||| |||
Db   843 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTRADGGGRMKCAACDGTCWAAWTCKAYTTNB 902

Qy   9066 AGTTTTTTTTTTTTTTTTTTTTTT 9090
            |||||||||||||||||||
Db   903 BTTTTTTTTTTTTTTTTWTTTT 927

RESULT 5
CNS005TE
LOCUS   CNS005TE   997 bp   DNA   linear   GSS 03-JUN-1999
DEFINITION   Drosophila melanogaster genome survey sequence TET3 end of BAC #
              BACR12K22 of RPCI-98 library from Drosophila melanogaster (fruit
              fly), genomic survey sequence.
ACCESSION   AL060767
VERSION   AL060767.1   GI:4943573
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
              Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
              Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
              Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 997)
AUTHORS   Genoscope.
TITLE   Direct Submission
JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
              collaboration with the Berkeley Drosophila Genome Project (BDGP).
              The BDGP is constructing a physical map of the Drosophila
              melanogaster genome using these BACs. For further information
              please see http://www.fruitfly.org The BDGP Drosophila
              melanogaster BAC library was prepared by Kazutoyo Osoegawa and
              Aaron Mammoser in Pieter de Jong's laboratory in the Department of
              Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
              NY. The library is named RPCI-98 and was constructed by partial
              EcoRI digestion of Drosophila DNA provided by the BDGP from the
              isogenic strain y2; cn bw sp; the same strain used for the BDGP's
              P1 and EST libraries. A more detailed description of the library,
              and how to order individual BAC clones, the entire library, or
              filters for hybridization from the BACPAC Resource Center can be
              found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
     source   1..997
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR12K22"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

Query Match   0.5%;   Score 48.4;   DB 9;   Length 997;
Best Local Similarity   16.2%;   Pred. No. 0.079;
Matches 85;   Conservative 189;   Mismatches 250;   Indels 0;   Gaps 0;

Qy   2239 AGCCTACCTCGTTACTTGTCCCTTTGTTTGGGGCGGCCTTCTGGTTACCCTTTGCGTCC 2298
            | :::| ||:| |::: | ::: ::|||::| |::::   ::::::Y:Y:Y:Y:Y:Y:Y:
Db   464 AYYYTCYCTYTYYYTTYTTTTYYTTYTTTYYYCCYYCCYYTCCYYYYYYYYYYYYTY 523

Qy   2299 TGTGCTCCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGTTTGTCTAAAGCTCAAGT 2358
            : :|: :||:|| |:: :| |::: |:   :::   :::|::|| :|:|| |||||
Db   524 YTYCTCYCTYCTCYTTYYYYTYYYYYYYYYYYYYYYTYYTTCYTTYYYCYCYYYYYYCYTCTCY 583

Qy   2359 AGCTCCTTTGCTTTGATTTTCTTCATCTGTTGCTATCTCCGCTGCAGGCTAGTTATGC 2418
            || |||| | |::: |||:|||::: :|||: |:::: |:|::|||:||: ::: :
Db   584 TYYYYCYCYTYYTCYCYYTYYTCYCYYYCYCYCYYYYYCYCYYYYYCYYYYYYTYYYYTT 643

Qy   2419 TGCCCTTTTAGGGTTTGTGTCCCATGGCCATGCGTCATTGTTTTTGCACC 9005
            |:||:||| :| ||| || : ||| ||| ||
Db   2419 TGCCCTTTTAGGGTTTGTGTCCCATGGCCATGCGCGGGCTTGCCCCTAACTTTCTTTTGTTGCAGC 2478

Db   471 CCYCTTGGTGCGTTTTTTTTTTTTKTTGTTTTTTTTTTTCTTTTTTCTTTTTTTTTTTTT 530

Qy   2738 TGTGCTTTAGTTAACTCCTATCTTGGCAACGTTGGGAGGAATTGGTTTTGGAACGTTACA 2797
                | || :||:|| ||||||| ||:: | ||||||| ||| ::|:: ::::
Db   531 KTCGGTKGKGTKTKTYTTTSTGTKTGKGTKTKTKKTKTKTSGGGKKKKTTYKBKTGT 590

Qy   2798 CTAAGACCGGAGAGTTTTCCTTGTCGTGGTTGTTCCCCGGTGCGACATATGACGCG 2857
                :|:| | | : ||| | ||||||| |||| | |||| :|||||||||||||
Db   591 GTSTGKGKGKGTKKTTTGTKKTTGKGTKTKKTKKKTKTTKGKKKGTGKGK 650

Qy   2858 CTGGTGACTTTCTGTGTGTCACGTAGCTCTTCTATGTTAACATCCAGTGCGAGCATCG 2917
                ||||| ||:|:|:| ||:||| :|:|: :|||: :|:: ::|:|
Db   651 KTGTTGGGGKKKTKTKKKKTKGKGKKGTKKTKTKTKGGKGKGKKKKKKKT 710

Qy   2918 TTCTTTGGGACTGACTCTAGGGTTAGGGCCCATAGAATGTTGGTGCGTCTCGGAAAGTGT 2977
                | | |:|:| |: |:||| | ::||| ::|: :::|| :|| :||
Db   711 KKTTKKTTKTKKKKKTKTKTKTKKGKKGGGKTKKGKGKKGKKGGTGTKKKG 770

Qy   2978 CATGCTTGGTATTCCATTATGTTTCTTAAGTTTTTCCTCTTAGTGTTTGGTGAGAATGGT 3037
                 ::| :: | ::::::||| :|||::|||||||::
Db   771 TGGKGKKGKGKKKKKGKKKKKKGKKKKKTTGKGTGKTKKTKKTKTKKKKKKKKKTKKKKGT 830

Qy   3038 GTGTTTTCT 3047
                ::| ||:|:|
Db   831 KKGGKTKKGT 840

RESULT 7

CNS006ST/c
LOCUS   CNS006ST   937 bp   DNA   linear   GSS   03-JUN-1999
DEFINITION   Drosophila melanogaster genome survey sequence TET3 end of BAC #
             BACR14F16 of RPCI-98 library from Drosophila melanogaster (fruit
             fly), genomic survey sequence.
ACCESSION   AL065880
VERSION   AL065880.1   GI:494848
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM   Drosophila melanogaster
             Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
             Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
             Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 937)
AUTHORS   Genoscope.
TITLE   Direct Submission
JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
             BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
             - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
             collaboration with the Berkeley Drosophila Genome Project (BDGP).
             The BDGP is constructing a physical map of the Drosophila
             melanogaster genome using these BACs. For further information
             please see http://www.fruitfly.org The BDGP Drosophila
             melanogaster genome using these BACs. For further information
             please see http://www.fruitfly.org The BDGP Drosophila
             melanogaster BAC library was prepared by Kazutoyo Osoegawa and
             Aaron Mammoser in Pieter de Jong's laboratory in the Department of
             Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
             NY. The library is named RPCI-98 and was constructed by partial
             EcoRI digestion of Drosophila DNA provided by the BDGP from the
             isogenic strain y2; cn bw sp, the same strain used for the BDGP's
             P1 and EST libraries. A more detailed description of the library
             and how to order individual BAC clones, the entire library, or
             filters for hybridization from the BACPAC Resource Center can be
             found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
     source   1..937
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR14F16"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

Db   644 CTCYYTYYYYYYYYAYKYCYYYYTCTCYYYCYCYCYYYTCTCYCYCYCYYYYTTCYCTY 703

Qy   2479 AGCTGCTGCCCAACCAGATTATGACTGTGGGCGCGACTGCTAGTGGCAGGGTTAGTTTT 2538
                :: ::|:| | | :   :   : :: | || ||:| :| ||:||
Db   704 TTYTYYYYYYTCTCYYYYYYYTTYTTTCTYYYYYYYYTTTYTYYTTTYTTTY 763

Qy   2539 GTGGGCCGGCCGCTAACCGTGGTCACCGCATAGCTCTGCTTGGAGTCCTTGGCCTCTGGT 2598
                :| : | ::| |: | :|::||||| | || || |:|| :|||
Db   764 YYCYYYCCYCYYYYCYCYYYYCTYYYYYYYYYYYTYTYYTYTYTTYTCTYYY 823

Qy   2599 AGCGCTTTAACCCTCTTGCATTGGTTACGGCCTGCTTCAGCTTTGATACCGAGATAAT 2658
                : : |::| ::| :: | |:|||| ::|| :||||: :|:| ::|:|| ||
Db   824 TYTYYYCYCYCYYCCCTCYTYYYCYCTYYYYYYCYTYTMTYYYYTTYYYTYTYTYYTYT 883

Qy   2659 TGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCATGTCTGTTTTGGCTTCTT 2718
                :   : |:| : | || |: | : |: |||: || :| |||| ::|:|
Db   884 YYTYHYTYTYTTTYYCCYCYCYCCCTSYCYCTYCYYCYTYYYYYTTYTYYTYCTCYYYTC 943

Qy   2719 TGCTCACTTGTTACCTCGCTGTGCTTTAGTTAACTCCTATCTTT 2762
                : |:|:|| ::: | ::|: |: |:|::| ||: :::|:
Db   944 TYTTYTYTYTYTCYTYYYYYYYYYYYTCYYYYYTYYYCYYYYTCYYYYTYTY 987

RESULT 6

CNS0073W
LOCUS   CNS0073W   922 bp   DNA   linear   GSS   03-JUN-1999
DEFINITION   Drosophila melanogaster genome survey sequence TET3 end of BAC #
             BACR14D09 of RPCI-98 library from Drosophila melanogaster (fruit
             fly), genomic survey sequence.
ACCESSION   AL066784
VERSION   AL066784.1   GI:4945247
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM   Drosophila melanogaster
             Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
             Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
             Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 922)
AUTHORS   Genoscope.
TITLE   Direct Submission
JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
             BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
             - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
             collaboration with the Berkeley Drosophila Genome Project (BDGP).
             The BDGP is constructing a physical map of the Drosophila
             melanogaster genome using these BACs. For further information
             please see http://www.fruitfly.org The BDGP Drosophila
             melanogaster BAC library was prepared by Kazutoyo Osoegawa and
             Aaron Mammoser in Pieter de Jong's laboratory in the Department of
             Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
             NY. The library is named RPCI-98 and was constructed by partial
             EcoRI digestion of Drosophila DNA provided by the BDGP from the
             isogenic strain y2; cn bw sp, the same strain used for the BDGP's
             P1 and EST libraries. A more detailed description of the library
             and how to order individual BAC clones, the entire library, or
             filters for hybridization from the BACPAC Resource Center can be
             found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
     source   1..922
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR14D09"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

Query Match   0.5%; Score 48; DB 9; Length 922;
Best Local Similarity 21.4%; Pred. No. 0.1;
Matches   79; Conservative 120; Mismatches 171; Indels 0; Gaps   0;

Qy   2678 CCTGTAGTAGCATTAGTTGTCATGTCTGTTTTGGCTTCTTTCGTTTGCTCACTTGTTACCTCGC 2737

Left column (RESULT 8)

```
Query Match         0.5%;  Score 47.4;  DB 9;  Length 937;
Best Local Similarity  23.0%;  Pred. No. 0.15;
Matches   86;  Conservative 133;  Mismatches 154;  Indels   1;  Gaps   1;

Qy  2693 GTTGTCATGTCGTTTGGCTTCTTTGCTCACTTGTTACCTCGCTGTCTT-TAGTTAA 2751
         | ::::::|::||:|| ::||||: ||:: ||:|:|| |:||:|||:||:|| :||||||
Db   932 GCBBYBBBKYCTBGBTBGBGTBSBSTTSGBTSGYKBKKKKKBBKTKTBGBKKCKBBTTBTTGB 873

Qy  2752 CTCCTATCTTTGGCAACGTCGGAGAATTGGTTTTGAACGTACACTAAGACCGGAGAG 2811
         :||| :::: ::||::|: |||:|:||:|:|||:
Db   872 KBKCTCGBSKKCKTBBKCKCGTNWDAAAAKATAKTDTTKTKKSBBGYGTSCKBCSKBBYTC 813

Qy  2812 GTTTTTCCTTGTGCTGGTTTGTTTTCCCCGGTGCGACATATGACGGCGCTGGTGACTTTTCTG 2871
         :||||:|||: :::||:||: ::|:::
Db   812 TSTTTYKTYSTBYBTBCKBTGYTKCTCSTSYKCCBBCSYSCYTSCSCYBBGTTCYYCYYS 753

Qy  2872 TGTGTGTCACGTAGCTCTTCTATGTTTAACATCCAGTGCGCAGCATCGTTCTTTGGGACTGA 2931
         |||:||: ||: :|||||: ||: |
Db   752 TSTCHSTTTYKGKTKTKTTBBYKYSTYTKGYTYKYKYKCBYKTYTKGBKKTKBKKYTTKG 693

Qy  2932 CTCTAGGGTTAGGGCCCATAGAATGTTGGTCGCGTCTCGGAAAGTGTCATGCTTGGTGATTC 2991
         :: ||||: :|: ||||||
Db   692 BGTGCKBKBTTTGKKKKKKKTGTKTTTTKKKTTTCTKGTTTTTKKKKTKKKTTKKTTGTGK 633

Qy  2992 TCATTATGTTCTTAAGTTTTCCCTCTTAGTGTTTGGTGAGAATGTGTGTTTTTCTATAA 3051
         :: ::|::: ::|||| ::||::: :|::::
Db   632 KTGKKGTKTTTKTTTKKTTKKTTTKTKTTTGKKKGTKKTKTTTTTGTKTTTTKTKT 573

Qy  3052 GCACTTGCATGGTG 3065
         |: | ||: |:||
Db   572 GTTTTKGKGTKTKK 559
```

RESULT 8
CNS006U0
LOCUS   CNS006U0   884 bp   DNA   linear   GSS 03-JUN-1999
DEFINITION   Drosophila melanogaster genome survey sequence T7 end of BAC #
                BACR14N21 of RPCI-98 library from Drosophila melanogaster (fruit
                fly), genomic survey sequence.
ACCESSION   AL065923
VERSION   AL065923.1   GI:4944891
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM   Drosophila melanogaster
                Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
                Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
                Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 884)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
                BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
                - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
                collaboration with the Berkeley Drosophila Genome Project (BDGP).
                The BDGP is constructing a physical map of the Drosophila
                melanogaster genome using these BACs. For further information
                please see http://www.fruitfly.org The BDGP Drosophila
                melanogaster BAC library was prepared by Kazutoyo Osoegawa and
                Aaron Mammoser in Pieter de Jong's laboratory in the Department of
                Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
                NY. The library is named RPCI-98 and was constructed by partial
                EcoRI digestion of Drosophila DNA provided by the BDGP from the
                isogenic strain y2; cn bw sp, the same strain used for the BDGP's
                P1 and EST libraries. A more detailed description of the library
                and how to order individual BAC clones, the entire library, or
                filters for hybridization from the BACPAC Resource Center can be
                found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
     source   1..884
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"

Right column (RESULT continuation and RESULT 9)

```
                                                     /clone="BACR14N21"
                                                     /clone_lib="RPCI-98"
                                                     /note="end : T7"

ORIGIN

Query Match         0.5%;  Score 46.2;  DB 9;  Length 884;
Best Local Similarity  16.4%;  Pred. No. 0.35;
Matches   44;  Conservative 114;  Mismatches 110;  Indels   0;  Gaps   0;

Qy  2806 GGAGAGGTTTTCCTTGTGCTGGTTTGTTTTCCCCGGTGCGACATATGACGGCGCTGGTGAC 2865
         ||||||| :|||||| ::|: |||: :||:
Db   596 GGKGKGGKTTTKTTKTKTKTKTTGKTGTGTTKKKGGGGGGGGKKKGKGGTGKG 655

Qy  2866 TTTCTGTGTGTCACGTAGCTCTTCTATGTTAACATCCAGTGCAGCATCGTTCTTTGG 2925
         : |:|:|| :||| |||||||||:|| :: |:|
Db   656 GGGGKKKKGGTKTKGGKDGTGGKKKTTKKKKKKSGTKGKGKGKGKKKKSGTKKKKSGGTGKG 715

Qy  2926 GACTGACCTCAGGGTTAGGGCCCATAGAATGTTGGTCGCGTCTCGGAAAGTGTCATGCTTG 2985
         :||:::||: :|:: : :|||:: |:
Db   716 TGTKGTGTKKKKGGGKGKGGKKKKKKGKKKKKGKKKKKTGKKTTKGKKKG 775

Qy  2986 GTATTCTCATTATGTTCTTAAGTTTTCCCTCTTAGTGTTTGGTGAGAATGTGTGTTTTTT 3045
         || ::: /:|::: ::::::::::::::::::::::
Db   776 GTKKKGKKGKKTKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKTKKKKKKKKKKK 835

Qy  3046 CTATAAGCACTTGCATGGTGATGTCTTG 3073
         |: ::: |:::: :::::::::::
Db   836 CKKKKKKKKKKKKKKKKKKKKKKKKKKNKBK 863
```

RESULT 9
CNS0071A
LOCUS   CNS0071A   895 bp   DNA   linear   GSS 03-JUN-1999
DEFINITION   Drosophila melanogaster genome survey sequence TET3 end of BAC #
                BACR14B09 of RPCI-98 library from Drosophila melanogaster (fruit
                fly), genomic survey sequence.
ACCESSION   AL066286
VERSION   AL066286.1   GI:4945153
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM   Drosophila melanogaster
                Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
                Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
                Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 895)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
                BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
                - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
                collaboration with the Berkeley Drosophila Genome Project (BDGP).
                The BDGP is constructing a physical map of the Drosophila
                melanogaster genome using these BACs. For further information
                please see http://www.fruitfly.org The BDGP Drosophila
                melanogaster BAC library was prepared by Kazutoyo Osoegawa and
                Aaron Mammoser in Pieter de Jong's laboratory in the Department of
                Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
                NY. The library is named RPCI-98 and was constructed by partial
                EcoRI digestion of Drosophila DNA provided by the BDGP from the
                isogenic strain y2; cn bw sp, the same strain used for the BDGP's
                P1 and EST libraries. A more detailed description of the library
                and how to order individual BAC clones, the entire library, or
                filters for hybridization from the BACPAC Resource Center can be
                found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
     source   1..895
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR14B09"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"

ORIGIN
```
                          /mol_type="genomic DNA"
                          /db_xref="taxon:7227"
                          /clone="BACR09023"
                          /clone_lib="RPCI-98"
                          /note="end : TET3"

Query Match     0.5%;  Score 46; DB 9; Length 947;
Best Local Similarity   36.8%;  Pred. No. 0.41;
Matches   93; Conservative  41; Mismatches 119; Indels   0; Gaps   0;

Qy  2188 GGCCGTGGTGCTAATTCTGTGGGGTTGTGTGGCAGCAAGTATCTTATTTAGCCTACCT 2247
         |: .||||:||::|||||||||||||||||||||::::::: ::||||||||
Db   592 GGGTKTTTKTTTKYTTTTTTGGTTTTTSCSBBGSYTCBYTTTTTTTTTTTTTTT 533

Qy  2248 CTGTTACTTGCTCCCTTGTTTGGGCGGCCTTCTGTGTGTTACCCTTGCGTCCTGTGCTCCC 2307
         |||||||||||||| :|||||||||||||||::|:||||::||
Db   532 TTTTTTTTTTTTGTSTNNNTGTNTNKTTTTTKTTTSYKTTTTTYKSCBTSTTTTKCC 473

Qy  2308 ATCCCAGTCGTATCTCCAAGCTGCTCGGGATGTTTTGTCTAAAGTCAAGTAGTCCTTT 2367
         :::::||||||||||||||||||||:::||||:::||
Db   472 GTYYGCGCGTCTTCGTTCGTTKTTTTTTTBTTTTTGGGTCTGKBTGGSTTTGGKGTTGC 413

Qy  2368 TGCTTGATTTCTTCATCTGTTGCTAATCTCCGCTGCAGGCTACGTTATGCTGCCCTTTT 2427
         ||::|||||:|:||| |:||| ||:|||||||||||||||:||
Db   412 TGYCTTTTTTTTTSSSCSCSTSTTTTGTCKYTTYYTTKKCKSKTGTTTTTTTTTBTSTTTTTBT 353

Qy  2428 AGGGTTTGTGCCC 2440
         ||||||||:||
Db   352 TTTTTTTTTBTC 340
```

RESULT 11
AG484307
```
LOCUS       AG484307              841 bp    DNA     linear   GSS 04-JUN-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-379M10.T7, genomic survey
            sequence.
ACCESSION   AG484307
VERSION     AG484307.1  GI:48191537
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     BAC end Sequences of Library MSMg01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 841)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : T7
            LIBRARY
            Vector    : pBACe3.6
            R.Site 1  : EcoRI
            R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1..841
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
```

ORIGIN
```
Query Match     0.5%;  Score 46; DB 9; Length 895;
Best Local Similarity   21.5%;  Pred. No. 0.4;
Matches   81; Conservative 128; Mismatches 163; Indels   4; Gaps   1;

Qy  2711 GGCTTCTTTGCTCACTTGTACCTGCGTGCTTAGTTAACTCCTATCTTTCGGCAACGT 2770
         :: |||: |||::|| |:||::  |: ||: |: ::|||:||
Db   481 KGKGKKTTTKKTYGGKTKKGGKGGGGKGKGKGKKKGSKGKKKTKKTKKTGGKKG 540

Qy  2771 TGGGAGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTTTCCTTGTGCTGTT 2830
         ||:|| ::|||||||||||:: : ||:  :|||||:||
Db   541 TGKGTKKKKGKKKTKTGSKKKKGTGTKTTKKKKTKGKTKKKTKGKTKKKKGTGKTGKKG 600

Qy  2831 TGTTTCCCCGGTGCGACATATGACGACGCGCTGGTGACTTTCGTGTGTGTCACGTAGCTCTT 2890
         ||:|| : ::|||:||||||:||||:|::||||||||
Db   601 TGKTKKTKTGKGKKKKGGGTGKKGTGGKGKKGGKTKTGKTGGKTKTKKKKKKTKKK 660

Qy  2891 CTATGTTTAACATCCAGTGCAGCATCGTTCTTTGGGACTGACTCTAGGGTTAGGGCCCAT 2950
         ||::|| ::||| |||:|:||||||:::
Db   661 KKKKGTGTKGGKKGKGKTG----KKKKKKGKTKTKKTTTTKTTTTKKTTGKKKTGGKG 716

Qy  2951 AGAATGTTGGTGCGTCTCGGAAAGTGTCATGCTTGGTATTCTCATTATGTTCTTAAGTTT 3010
         ||||| :|||:|::: |:||:||:||
Db   717 GGKGKGKTKGKKKGKGGGKGGGKKKKTKTKKKTKTKTGKGKKKKTKTTKKKKTGKTKKKKGTKKKK 776

Qy  3011 TTCCTCCTTAGTCGTTGTTGGTGAGAATGGTGTGTTTTTCTATAAGCCACTTGCATGGTGATGTC 3070
         ||: ::|||||:||| ::::|||:: :|||||:::
Db   777 TGTGKGKTKKKGKTGKGKKGTKGKGKGGKKKKKKKKTGKKTTTGTKGGKKKGGGKKTKKKKKKGKKT 836

Qy  3071 TTGCCTAATGATTTTG 3086
         ||:::|:::::||
Db   837 TKKKKGTKTKKKTKK 852
```

RESULT 10
CNS0047J/c
```
LOCUS       CNS0047J              947 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR09023 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL066360
VERSION     AL066360.1  GI:4943005
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 947)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp; the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1..947
                     /organism="Drosophila melanogaster"
```

/clone="BACR14J09"
/clone_lib="RPCI-98"
/note="end : T7"

ORIGIN

Query Match　　　0.5%;　Score 45;　DB 9;　Pred. No. 0.82;　Length 919;
Best Local Similarity　21.9%;
Matches　93;　Conservative　150;　Mismatches 178;　Indels　4;　Gaps　1;

```
Qy   2173 GGGAGTACCACCAAGCCGTGGTGCTAATTCTGTTGGGGTTGTGTGGCAGCAAGTATCT 2232
          ||:||::: .::       ::||||
Db    493 GGRTSTMHMTMYSTCGGTGWGYKYKTYTYCTTGTTGTKCCKTGYYTYTYTYSTSTY 552

Qy   2233 TATTTAGCCTACCTCTGTTACTTGTCCCTTGTTGTTTGGGGCGGCCTTCTGGTTACCCTT 2292
          |:|||:  :||:|||||||:   ||||||:||||
Db    553 TYYTTHTYCTTYCTCTCCYTCCTTKCTKTSTGYKTTTSTGCKKTGCTCGGCTYTKKTT 612

Qy   2293 GCGTCCTGTGCTCCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGTTTGTCTAAAGC 2352
          :::|| : :::||  ||:||       :::||
Db    613 BTKTTSBTSYGGYCCTCGYTYGTCTYCTTSCYBCBCGTSKKGGSYGGBBTSGY---GG 668

Qy   2353 TCAAGTAGCTCCTTTGCTTGATTTCTTCATCTGTTGCTATCTCCGCTGCAGGCTACG 2412
          :|:::|||:::|||:||  |||::|||:     ::||::
Db    669 KCKKGBTTCTGBTSKCTGBTGYSKCTGKKCYTYSTYTCTTKKTYTCTTYTYYTYYST 728

Qy   2413 TTATGCTGCCCTTTAGGGTTTGTGCCCATGGCTGCGGGCTTGCCCCTAACTTTCTTTGT 2472
          |||:::||||  |:|:| |||:||      :|:::|: ::|||::
Db    729 TTTTSTYBTTYKTTKTKTTKYTYGTCTGKCTCBTSYKKBKYGTBKBSKKTBTSYYGB 788

Qy   2473 TGCAGCAGTCTGCTGCCCAACCAGATTATGACTGTGGGTGCGGACTGCTAGTGCAGGGTT 2532
          ::::::||:|::: : .::  :|: :: ::|:::
Db    789 BSYSSKSTBVSBSSTBSSYSBTBBSSSTSSSSTSTBTSSBSSTSSSBBTSBS 848

Qy   2533 AGTTTTGTGGGCCGGCCTAACCGTGTCACCGCATAGCTCTGCTTGTAGGTCCTTGGCC 2592
          :::|:|||: ::::::: :|:|||::|:: :::::::::||
Db    849 SSSTSTSTSTBSTSBSSTSSSBBGSBBCTBSBBSSTSSSBBSTSTTSTSBBBSBSTSSS 908

Qy   2593 TCTGG 2597
          |:|||
Db    909 TBBSB 913
```

RESULT 13
CNS0164Y
LOCUS　　　　　CNS0164Y　　　　　　1201 bp　　　DNA　　　linear　　GSS 26-JUL-1999
DEFINITION　Drosophila melanogaster genome survey sequence T7 end of BAC
　　　　　　　BACN15M04 of DrosBAC library from Drosophila melanogaster (fruit
　　　　　　　fly), genomic survey sequence.
ACCESSION　 AL106300
VERSION　　　AL106300.1　GI:5621234
KEYWORDS　　GSS.
SOURCE　　　Drosophila melanogaster (fruit fly)
　ORGANISM　Drosophila melanogaster
　　　　　　Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
　　　　　　Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
　　　　　　Ephydroidea; Drosophilidae; Drosophila.
REFERENCE　 1　(bases 1 to 1201)
　AUTHORS　　Genoscope.
　TITLE　　　Direct Submission
　JOURNAL　　Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
　　　　　　 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
　　　　　　 - Web : www.genoscope.cns.fr)
COMMENT　　 Determination of this BAC-end sequence was carried out as part of a
　　　　　　collaboration with the European Drosophila Genome Project (EDGP) -
　　　　　　http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
　　　　　　library (Dros BAC) was made by Alain Billaud at CEPH (Centre
　　　　　　d'Etude du Polymorphisme Humain) with funding provided by a MRC
　　　　　　project grant. The DNA was prepared from embryos by Alain Bucheton
　　　　　　and Geneviéve Payan. It has been constructed in the vector
　　　　　　pBeloBAC11.
FEATURES　　　　　　Location/Qualifiers
　　source　　　　　1..1201
　　　　　　　　　　/organism="Drosophila melanogaster"

/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-379M10.T7"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN

Query Match　　　0.5%;　Score 45.4;　DB 9;　Length 841;
Best Local Similarity　51.8%;　Pred. No. 0.6;
Matches　127;　Conservative　0;　Mismatches 116;　Indels　2;　Gaps　1;

```
Qy   2226 AGTATCTTATTTAGCCTACCTCTGTTACTTGTCCCTTGTTTGGGCGGCCTTCTGGTT 2285
          |||||||: ||||:|||||||||||||||||: |||||||  ||||||||||||||||
Db    512 ACTTTCTCTTTTCCTTCTCCYTTTCCCCCTCTTTTTTTTCCCCCTCTTCTTTTT 571

Qy   2286 ACCCTTTGGCGTCCTGTGCTCCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGTTTGT 2345
          |||||||| |||| ||||||||||   |||||| ||| | |||||||||||||||||||||
Db    572 TTCCCT--CCTTTTTCTCCCCTGGCGATCCTTTTTTACCCTCCCTGCTCCCCTCTTTTTT 629

Qy   2346 CTAAAGCTCAAGTAGCTCCTTTGCTTGATTTCTTCATCTGTTGCTATCTCCGCTGCA 2405
          ||||||| |||||||||||||| || ||| ||||||||||||||||
Db    630 TTACCTTTCCCCTCTTTTTTTCATATCTTTTTTCCCCCTTTTCCTTATTCCATCCTT 689

Qy   2406 GGCTACGTTATGCTGCCCTTTAGGGTTTGTGCCCATGGCTGCGGGCTTGCCCCTAACTT 2465
          ||||||| |||||||||||| |||||||||||| |||||||||||||||||||||||||
Db    690 CTCTTCACCACTTACACCTTTCTTTGTTTCTTTTTTGTGTCTCTCCACTATGTT 749

Qy   2466 TCTTT 2470
          |||||
Db    750 TCTTT 754
```

RESULT 12
CNS006S5
LOCUS　　　　　CNS006S5　　　　　919 bp　　　DNA　　　linear　　GSS 03-JUN-1999
DEFINITION　Drosophila melanogaster genome survey sequence T7 end of BAC #
　　　　　　BACR14J09 of RPCI-98 library from Drosophila melanogaster (fruit
　　　　　　fly), genomic survey sequence.
ACCESSION　 AL065856
VERSION　　　AL065856.1　GI:4944824
KEYWORDS　　GSS.
SOURCE　　　Drosophila melanogaster (fruit fly)
　ORGANISM　Drosophila melanogaster
　　　　　　Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
　　　　　　Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
　　　　　　Ephydroidea; Drosophilidae; Drosophila.
REFERENCE　 1　(bases 1 to 919)
　AUTHORS　　Genoscope.
　TITLE　　　Direct Submission
　JOURNAL　　Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
　　　　　　 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
　　　　　　 - Web : www.genoscope.cns.fr)
COMMENT　　 Determination of this BAC-end sequence was carried out as part of a
　　　　　　collaboration with the Berkeley Drosophila Genome Project (BDGP).
　　　　　　The BDGP is constructing a physical map of the Drosophila
　　　　　　melanogaster genome using these BACs. For further information
　　　　　　please see http://www.fruitfly.org The BDGP Drosophila
　　　　　　melanogaster BAC library was prepared by Kazutoyo Osoegawa and
　　　　　　Aaron Mammoser in Pieter de Jong's laboratory in the Department of
　　　　　　Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
　　　　　　NY. The library is named RPCI-98 and was constructed by partial
　　　　　　EcoRI digestion of Drosophila DNA provided by the BDGP from the
　　　　　　isogenic strain y2; cn bw sp, the same strain used for the BDGP's
　　　　　　P1 and EST libraries. A more detailed description of the library
　　　　　　and how to order individual BAC clones, the entire library, or
　　　　　　filters for hybridization from the BACPAC Resource Center can be
　　　　　　found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES　　　　　　Location/Qualifiers
　　source　　　　　1..919
　　　　　　　　　　/organism="Drosophila melanogaster"
　　　　　　　　　　/mol_type="genomic DNA"
　　　　　　　　　　/db_xref="taxon:7227"

```
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN15M04"
/clone_lib="DrosBAC"
/plasmid="pBeloBAC11"
/note="end : T7"
ORIGIN

Query Match      0.5%;  Score 45;  DB 9;  Length 1201;
Best Local Similarity   41.6%;  Pred. No. 0.91;
Matches   57;  Conservative  33;  Mismatches  47;  Indels   0;  Gaps   0;

Qy   8954  AGAAGATTGCAGAAGTTCCTTGTGAAGTATTGGCTGTCATTGTTTTTGCCCTAGGGCTC  9013
                 |||||  |  ||| | |    |||||:|         | ||:||||    |||
Db    671  AGAAGAWCTKTWTGTTGCCAGNGCTATATKTDDWDGRATATWTWWHYNHYHHYHHYYY   730

Qy   9014  ATTGCTGTTGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAACTAACAGTTTTTT  9073
                 ||:.|        ::  |||       |||| || ||:|:||||
Db    731  YYTYYBGKCTAATWTHYHAAARAAWTAACTWGRWWDAAAAAAATADTTYKTTT        790

Qy   9074  TTTTTTTTTTTTTTT  9090
                 ||:||||||||||
Db    791  TTWKTTTTTTTTTTTT  807

RESULT 14
CNS0091P/c
LOCUS      CNS0091P        925 bp    DNA    linear    GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
           BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
           fly), genomic survey sequence.
ACCESSION  AL053013
VERSION    AL053013.1  GI:4934461
KEYWORDS   GSS.
SOURCE     Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
           Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
           Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
           Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1  (bases 1 to 925)
AUTHORS    Genoscope.
TITLE      Direct Submission
JOURNAL    Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
           BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
           - Web : www.genoscope.cns.fr)
COMMENT    Determination of this BAC-end sequence was carried out as part of a
           collaboration with the Berkeley Drosophila Genome Project (BDGP).
           The BDGP is constructing a physical map of the Drosophila
           melanogaster genome using these BACs. For further information
           please see http://www.fruitfly.org The BDGP Drosophila
           melanogaster BAC library was prepared by Kazutoyo Osoegawa and
           Aaron Mammoser in Pieter de Jong's laboratory in the Department of
           Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
           NY. The library is named RPCI-98 and was constructed by partial
           EcoRI digestion of Drosophila DNA provided by the BDGP from the
           isogenic strain y2; cn bw sp, the same strain used for the BDGP's
           P1 and EST libraries. A more detailed description of the library
           and how to order individual BAC clones, the entire library, or
           filters for hybridization from the BACPAC Resource Center can be
           found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
source     1..925
           /organism="Drosophila melanogaster"
           /mol_type="genomic DNA"
           /db_xref="taxon:7227"
           /clone="BACR19D16"
           /clone_lib="RPCI-98"
           /note="end : TET3"
ORIGIN

Query Match      0.5%;  Score 43.8;  DB 9;  Pred. No. 1.9;
Best Local Similarity   14.6%;
Matches   66;  Conservative 164;  Mismatches  222;  Indels   0;  Gaps   0;
```

```
Qy    40  CTCCTGGACACCCCCTAGCAGGGCGTGGGGGATTCCCCTGCCCCTCTGCGAGAAGGG     99
                ::: : :  :: : : :    :  :  ::: :: : :  :   :::: ::: :::::
Db   924  SBSCSCSCCSCSSBBCSSSSMSTSSSNSSBCCCCSCCSSBBSSSTSMSSSBSSSGSSSSSS  865

Qy   100  TGGAGCCAACCACCTTAGTATGTAGGCGCGGGGACTCATGACGCTCGCTGATGACAAGC   159
                :::: ::::    :  ::: ::  :::::: :  |::|::||  | ||:: ::||
Db   864  SGTTSSACVKCNASSSCGCCGCGMABCCMCSSSSSCCGSASARGCVKVRASGGAKGRGGSG  805

Qy   160  GCCAAGCTTGACTTGGATGGCCCTGATGGGCGTTCATGGGTTCGTGGTGGTGGCGCTTT   219
                 | |:|   || | : || :|||:||| | |:|| |  :|||:||   ||||
Db   804  GASASHSSSSACBSSSSSCSASCWSASSSSASSSRSRSSGGGAGGSGASSSRSSSSSSSA   745

Qy   220  AGGCAGCCTCCACGCCCACCACCTCCAGATAGAGCGGCGCGCACTGTAGGGAAGACCGGG   279
                ::::: ::  :::   ::  ::  :::    :  :  :  ::::::: ::||:||
Db   744  SAGGVVSSASSSSSSCSSSVSCSVASSMSCSSBSSSVSSAGGSSSSSASCASCSCCCT    685

Qy   280  GACCGGTCACTACCACCAGGACGCAGACCTCTTTTTGAGTATCAGCCTCCGGAAGTAGTTG   339
                ::: :|     | :  : |       :: ||   |||   | | :|| :| |||:||
Db   684  SWSCSCSTSASMSAARSSSSCSSSSMSASSSSASSSSASSSSSGSSSSSSGSSSACGBS   625

Qy   340  GGCAAGCCCCACCTATATGTGTTGGGATGGTTGGGGTTAGCCATCATACCGTACTGCCTG   399
                 :: : ::: ::: :  :|   :||| : ||  :| || |::: |  ::::::
Db   624  MSSCGGGSGSVSASSGMSSSVSSSGGRSSGSGGGSSGGSGVGGSGSSSSSGSGSGVCSCSS  565

Qy   400  ATAGGGTCCTTGCGAGGGGATCTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTC   459
                :  :::  :  :::::: : :  :||  : | :| ::|:|| |:||| |:|:||::|
Db   564  GCMCRCSCSSAAAAAASCVAASCGMMCGXKSGCTNNNNNTNTNTNTNTNNNNNTTAA    505

Qy   460  TACTCAAACAAGTCCTGTACCTGCGCCCAGAA   491
                || || ||:|| |||  ::::||
Db   504  TAAATAANNNTNTATTNNNNNTNNNNNNANAA   473

RESULT 15
CNS006QP
LOCUS      CNS006QP        902 bp    DNA    linear    GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
           BACR14J09 of RPCI-98 library from Drosophila melanogaster (fruit
           fly), genomic survey sequence.
ACCESSION  AL065804
VERSION    AL065804.1  GI:4944772
KEYWORDS   GSS.
SOURCE     Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
           Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
           Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
           Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1  (bases 1 to 902)
AUTHORS    Genoscope.
TITLE      Direct Submission
JOURNAL    Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
           BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
           - Web : www.genoscope.cns.fr)
COMMENT    Determination of this BAC-end sequence was carried out as part of a
           collaboration with the Berkeley Drosophila Genome Project (BDGP).
           The BDGP is constructing a physical map of the Drosophila
           melanogaster genome using these BACs. For further information
           please see http://www.fruitfly.org The BDGP Drosophila
           melanogaster BAC library was prepared by Kazutoyo Osoegawa and
           Aaron Mammoser in Pieter de Jong's laboratory in the Department of
           Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
           NY. The library is named RPCI-98 and was constructed by partial
           EcoRI digestion of Drosophila DNA provided by the BDGP's
           isogenic strain y2; cn bw sp, the same strain used for the BDGP's
           P1 and EST libraries. A more detailed description of the library
           and how to order individual BAC clones, the entire library, or
           filters for hybridization from the BACPAC Resource Center can be
           found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
source     1..902
           /organism="Drosophila melanogaster"
           /mol_type="genomic DNA"
```

```
                    /db_xref="taxon:7227"
                    /clone="BACR14J09"
                    /clone_lib="RPCI-98"
                    /note="end : TET3"

ORIGIN

Query Match         0.5%;   Score 43.4;  DB 9;   Length 902;
Best Local Similarity  15.1%;  Pred. No. 2.5;
Matches  42;  Conservative  140;  Mismatches  97;  Indels  0;  Gaps  0;

Qy   2247  TCTGTTACTTGTCCCTTGTTTGGGGCGGCTTCTGGTTACCCTTTGCGTCCTGTGCTCC 2306
                 |::::|  ||::   ::|:||||:::::   ::|:   :|:   :   |   :::::|
Db    620  TKTKKKTKTTTKKTKKTKTKKTKKGTTKYTKKGBKTKKTKBTCCKKKTTTBKKBKTTT  679

Qy   2307  CATCCCAGTCGTATCTCCAAGCTGGCTGGGATGTTTTGTCTAAAGCTCAAGTAGCTCCTT 2366
                 |   |      |:|   ||: ::|| :::: ::|:  : :|:    :|  ::::::|:|
Db    680  TKTTKTKTKYKTKTGBBTYCSGKTKBBKKKKKKKGKBGGGKGKGKGTBKBTKKKKKYTBT  739

Qy   2367  TTGCTTTGATTTCTTCCATCTGTGCTATCTCCGCTGCAGGCTACGTTATGCTGCCCTTT 2426
                 :::::|:::  :|:::::|  |:|| |:|::|||  ::::: |:::| :::: :::
Db    740  BBKYBTBBYKTKYBTTKTTTTBTTBCTTBYBTTTKYTYBBTBBBBKYKTBKSBKKBKY  799

Qy   2427  TAGGGTTTGTGCCCATGGCTGCGGGCTTGCCCCTAACTTCTTTTGTTGCAGCAGCTGCTG 2486
                 |::::::  ::: :|:: ::: :|:|::|    :::||  ::   :|: |: :::
Db    800  TBYSKBBYBYBTTYBTSYSSTBBBBTYKBTBSBYCCTBYTBYBBBSTCBBBSBSGSBSYSS  859

Qy   2487  CCCAACCAGATTATGACTGGTGGGTGCGACTGCTAGTGG 2525
                 :::  :|::   ::|| :::::::::::   :|::
Db    860  SSSYBBSBGBGTSTBKKKBKKTBYSBKBKBBBKKGBTBB  898


Search completed: October 30, 2005, 23:11:36
Job time : 17787 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2005, 06:47:36 ; Search time 1443 Seconds
(without alignments)
9609.733 Million cell updates/sec

Perfect score: 1679
Sequence: 1 gttgtgtcctcagcaaac........ataaaaagagcaaaaaa 1679

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9784742 seqs, 4129495052 residues

Total number of hits satisfying chosen parameters: 19569484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1500 summaries

Database : Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq2:*
25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Score	Length	DB	ID		Description
52	1679	100.0	1679	15	US-10-066-269-103		Sequence 103, App
64	1679	100.0	1679	15	US-10-167-749-522		Sequence 522, App
226	1679	100.0	1679	16	US-10-223-085-55		Sequence 55, Appl
232	1679	100.0	1679	16	US-10-219-065-125		Sequence 125, Appl
262	1679	100.0	1679	16	US-10-223-084-55		Sequence 55, Appl
263	1679	100.0	1679	16	US-10-223-088-55		Sequence 55, Appl

264	1679	100.0	1679	16	US-10-223-090-55		Sequence 55, Appl
269	1679	100.0	1679	16	US-10-223-087-55		Sequence 55, Appl
271	1679	100.0	1679	16	US-10-223-083-55		Sequence 55, Appl
274	1679	100.0	1679	16	US-10-223-089-55		Sequence 55, Appl
449	1679	100.0	1679	17	US-10-223-081-55		Sequence 55, Appl
487	1679	100.0	1679	18	US-10-170-481A-522		Sequence 522, App
556	1679	100.0	1679	18	US-10-162-521A-522		Sequence 522, App
558	1679	100.0	1679	18	US-10-162-521A-522		Sequence 522, App
618	1679	100.0	1679	18	US-10-305-654-55		Sequence 55, Appl
626	1679	100.0	1679	19	US-10-081-056-55		Sequence 55, Appl
641	1679	100.0	1679	22	US-10-918-851-522		Sequence 522, App
660	1679	100.0	1679	22	US-10-931-886-375		Sequence 375, App
661	1679	100.0	1679	22	US-10-805-667-522		Sequence 522, App
662	1679	100.0	1679	22	US-10-897-355-522		Sequence 522, App
663	1679	100.0	1679	22	US-10-893-802-522		Sequence 522, App
665	1679	100.0	1679	24	US-10-897-360-522		Sequence 522, App
666	1679	100.0	1679	24	US-10-995-952-375		Sequence 375, App
667	1679	100.0	1679	26	US-11-129-762-522		Sequence 522, App
670	1693	99.0	1661.4	19	US-11-129-762-522		Sequence 2, Appli
671	2012	97.9	1643.4	9	US-09-966-546-5		Sequence 5, Appli
672	2012	97.9	1643.4	9	US-09-966-545-5		Sequence 5, Appli
673	2012	97.9	1643.4	10	US-09-965-212-5		Sequence 5, Appli
674	2012	97.9	1643.4	24	US-10-971-479-5		Sequence 5, Appli
675	1603	95.4	1601.4	9	US-09-966-546-3		Sequence 3, Appli
676	1603	95.4	1601.4	9	US-09-966-545-3		Sequence 3, Appli
677	1603	95.4	1601.4	10	US-09-965-212-3		Sequence 3, Appli
678	1603	95.4	1601.4	16	US-10-189-940-3		Sequence 3, Appli
679	1603	95.4	1601.4	24	US-10-971-479-3		Sequence 3, Appli
680	1678	86.8	1457	14	US-10-098-841-72		Sequence 72, Appl
681	3987	85.7	1438.4	21	US-10-723-860-5804		Sequence 5804, Ap
682	3987	85.7	1438.4	21	US-10-723-860-7421		Sequence 7421, Ap
683	1839	85.3	1432.8	15	US-10-161-572-16		Sequence 16, Appl
684	1839	85.3	1432.8	18	US-10-295-027-455		Sequence 455, App
685	2129	84.5	1418.2	16	US-10-306-013-2		Sequence 2, Appli
686	1032	61.5	1032	19	US-10-657-103-1		Sequence 1, Appli
687	1061	58.0	974	15	US-10-161-572-19		Sequence 19, Appl
688	939	55.9	931	19	US-10-657-105-5		Sequence 5, Appli
689	1094	55.4	891.8	15	US-10-161-572-17		Sequence 17, Appl
690	1130	53.1	539.2	18	US-10-161-572-18		Sequence 18, Appl
691	1478	32.1	539.2	35	US-10-161-572-35		Sequence 35, Appl
692	3092	32.1	539.2	22	US-10-696-639-8		Sequence 8, Appli
693	3110	32.1	537.6	15	US-10-161-572-32		Sequence 32, Appl
694	1071	32.0		15	US-10-161-572-34		Sequence 34, Appl
695	1080	32.0		15	US-10-161-572-33		Sequence 33, Appl
696	503	30.0	503	18	US-10-167-749-524		Sequence 524, App
697	503	30.0	503	18	US-10-170-481A-524		Sequence 524, App
732	503	30.0	503	18	US-10-210-028-524		Sequence 524, App
768	503	30.0	503	18	US-10-162-521A-524		Sequence 524, App
770	503	30.0	503	18	US-10-805-667-524		Sequence 524, App
775	503	30.0	503	22	US-10-897-359-524		Sequence 524, App
787	503	30.0	503	22	US-10-893-802-524		Sequence 524, App
788	503	30.0	503	22	US-10-897-360-524		Sequence 524, App
789	503	30.0	503	26	US-11-129-762-524		Sequence 524, App
790	537	29.3	503	18	US-10-378-029-21		Sequence 21, Appl
791	617	26.8	491.2	18	US-10-187-975-107		Sequence 107, App
793	452	26.3	449.4	15	US-10-052-281-167		Sequence 167, App
794	437	24.1	441.2	10	US-09-918-995-2450		Sequence 2450, Ap
795	484	23.6	404	24	US-10-450-763-13839		Sequence 13839, A
796	591	22.1	396	13	US-09-925-065A-298075		Sequence 298075,
797	409	21.1	370.6	10	US-09-918-995-33488		Sequence 33488, A
798	1195	20.3	353.8	14	US-10-044-090-560		Sequence 560, App
799	1017	20.3	341.6	15	US-10-161-572-28		Sequence 28, Appl
800	1017	20.1	340.6	15	US-10-161-572-30		Sequence 30, Appl
801	1757	20.1	340.6	10	US-09-978-418-37		Sequence 37, Appl
802	1757	20.0	337.4	22	US-10-485-231-37		Sequence 37, Appl
803	1075	19.4	337.4	15	US-10-161-572-29		Sequence 29, Appl
804	1153	18.7	335	16	US-10-106-698-1517		Sequence 1517, Ap
805	898	17.7	325.2	15	US-10-161-572-31		Sequence 31, Appl
806	443	16.6	314.6	24	US-10-450-763-13838		Sequence 13838, A
807	1809		296.8	15	US-10-161-572-27		Sequence 27, Appl

c

Left table — columns: No. | Score | % | Accession | Len | Pos | Sequence

No.	Score	%	Accession	Len	Pos	Sequence
811	272.6	16.2	US-10-038-854-25	18	1165	Sequence 25, Appl
812	272.6	16.2	US-10-038-854-23	18	1196	Sequence 23, Appl
813	272.6	16.2	US-10-312-352-51	19	1327	Sequence 51, Appl
820	272.6	16.2	US-10-066-269-129	15	4834	Sequence 129, App
868	272.6	16.1	US-10-219-065-1	16	4834	Sequence 1, Appl
928	271	16.1	US-10-042-865-13	18	1119	Sequence 13, Appl
929	271	16.1	US-10-104-047-995	18	2383	Sequence 995, App
964	269.4	16.0	US-10-167-749-611	15	2840	Sequence 611, App
1000	269.4	16.0	US-10-170-481A-611	18	2840	Sequence 611, App
1002	269.4	16.0	US-10-210-028-611	18	2840	Sequence 611, App
1007	269.4	16.0	US-10-162-521A-611	18	2840	Sequence 611, App
1019	269.4	16.0	US-10-918-851-611	18	2840	Sequence 611, App
1020	269.4	16.0	US-10-805-667-611	22	2840	Sequence 611, App
1021	269.4	16.0	US-10-897-359-611	22	2840	Sequence 611, App
1022	269.4	16.0	US-10-893-802-611	22	2840	Sequence 611, App
1023	269.4	16.0	US-10-897-360-611	22	2840	Sequence 611, App
1024	269.4	16.0	US-11-129-762-611	26	2840	Sequence 611, App
1025	269.4	16.0	US-10-015-115-23	26	1017	Sequence 23, Appl
1026	260.4	15.5	US-10-471-449-21	19	1427	Sequence 21, Appl
1027	260.4	15.5	US-10-480-172-7	20	2653	Sequence 7, Appli
1028	260.4	15.4	US-10-015-115-21	18	1018	Sequence 21, Appl
1030	258.8	15.4	US-10-015-115-25	18	1136	Sequence 25, Appl
1031	257.2	15.3	US-10-332-947-25	19	1011	Sequence 25, Appl
1032	246	14.7	US-10-015-115-27	19	1169	Sequence 27, Appl
1033	246	14.7	US-09-764-853-141	9	2813	Sequence 141, App
1034	240.8	14.3	US-10-480-172-13	20	2601	Sequence 13, Appl
1035	234	13.9	US-10-332-947-24	19	1056	Sequence 24, Appl
1036	232	13.8	US-10-091-438-57	15	4656	Sequence 57, Appl
1037	220.6	13.1	US-10-450-763-13807	24	2883	Sequence 13807, A
1038	217.8	13.1	US-10-480-172-9	20	754	Sequence 9, Appli
1039	211.4	12.6	US-10-480-172-11	19	2597	Sequence 11, Appl
1040	208.6	12.4	US-09-925-065A-544756	13	620	Sequence 544756,
1041	173.6	10.3	US-10-450-763-13841	24	767	Sequence 13841, A
1042	170.2	10.1	US-09-925-065A-156319	13	169	Sequence 156319,
1043	168.6	10.0	US-09-925-065A-156318	13	169	Sequence 156318,
1044	167	9.9	US-09-925-065A-72637	24	585	Sequence 72637, A
1045	163.2	9.7	US-10-052-283-418	15	600	Sequence 418, App
1046	160	9.5	US-10-016-253-8	16	408	Sequence 8, Appli
1047	160	9.5	US-10-991-173-8	16	408	Sequence 8, Appli
1048	156	9.3	US-10-027-632-12436	14	619	Sequence 12436, A
1049	156	9.3	US-10-027-632-12436	18	619	Sequence 12436, A
1050	143.4	8.5	US-09-925-065A-294135	13	537	Sequence 294135,
1051	137.2	8.2	US-10-450-763-13842	24	693	Sequence 13842, A
1052	121	7.2	US-09-925-065A-77258	13	601	Sequence 77258, A
1053	110.8	6.6	US-09-908-975-13736	10	125	Sequence 13736,
1054	107	6.4	US-10-723-860-4801	21	1275	Sequence 4801, App
1055	107	6.4	US-10-450-766-7708	24	1275	Sequence 7708, App
1056	107	6.4	US-10-450-763-13840	24	1275	Sequence 13840, A
1057	105.6	6.3	US-10-450-763-14396	24	913	Sequence 14396, A
1058	94.4	5.6	US-10-120-988-410	18	2678	Sequence 410, App
1059	94.4	5.6	US-10-450-763-3050	24	2678	Sequence 3050, App
1060	94.4	5.6	US-10-450-763-7527	24	3131	Sequence 7527, App
1061	89	5.3	US-10-450-763-249	24	3131	Sequence 249, App
1062	76.8	4.6	US-10-450-763-602	24	2026	Sequence 602, Appl
1063	73	4.3	US-10-676-248B-6	13	566	Sequence 6, Appli
1064	73	4.3	US-09-764-853-347	9	293	Sequence 347, App
1065	63.8	3.8	US-10-091-438-112	15	293	Sequence 112, App
1066	63.8	3.8	US-10-424-599-31621	19	1392	Sequence 31621, A
1067	61.2	3.6	US-10-027-632-17789	14	718	Sequence 17789, A
1068	61.2	3.6	US-10-027-632-17789	18	718	Sequence 17789, A
1069	60	3.6	US-09-908-975-13736	10	60	Sequence 13736,
1070	57.2	3.4	US-10-021-323-9953	20	248	Sequence 9953, App
1071	56.6	3.4	US-09-960-352-12773	20	399	Sequence 12773, A
1072	55.8	3.3	US-10-767-701-20328	20	665	Sequence 20328, A
1073	55.6	3.3	US-10-357-930-56798	20	247	Sequence 56798, A
1074	55.4	3.3	US-10-767-701-14981	21	3569	Sequence 14981, A
1075	55	3.3	US-10-425-115-140282	21	440	Sequence 140282,
1076	55	3.3	US-10-425-115-96437	21	563	Sequence 96437, A
1077	55	3.3	US-10-257-166-26	18	6359	Sequence 26, Appl
1078	54.4	3.3	US-10-425-115-36465	19	1199	Sequence 36465, A
1079	54.4	3.2	US-09-918-995-2447	10	514	Sequence 2447, App
1080	54.4	3.2	US-10-311-455-761	16	6171	Sequence 761, Appl
1081	53.8	3.2	US-10-425-115-14467	21	489	Sequence 14467, A

Right table — columns: No. | Score | % | Pos | Len | Accession | Sequence

No.	Score	%	Pos	Len	Accession	Sequence
1082	53.6	3.2	731	21	US-10-425-115-169207	Sequence 169207
1083	53.4	3.2	410	20	US-10-435-115-105926	Sequence 105926
1084	53.4	3.2	1189	21	US-10-437-963-40780	Sequence 40780, A
1085	53.4	3.2	1971	21	US-10-425-115-106633	Sequence 106633
1086	53	3.2	607	21	US-10-357-930-58712	Sequence 58712, A
1087	53	3.2	663	21	US-10-425-115-120879	Sequence 120879,
1088	53	3.2	1171	21	US-10-425-115-14321	Sequence 14321, A
1089	52.8	3.1	282	21	US-10-425-115-141273	Sequence 141273,
1090	52.8	3.1	3351	17	US-10-134-102-5	Sequence 5, Appli
c1091	52.8	3.1	5567	16	US-10-311-455-1562	Sequence 1562, Ap
c1092	52.6	3.1	5567	19	US-10-240-589C-82	Sequence 82, Appl
1093	52.6	3.1	2273	21	US-10-425-115-14026	Sequence 14026, A
c1094	52.4	3.1	579	21	US-10-363-345A-12021	Sequence 12021, A
1095	52.4	3.1	579	21	US-10-363-345A-12022	Sequence 12022, A
c1096	52.4	3.1	579	22	US-10-363-483A-12021	Sequence 12021, A
1097	52.4	3.1	579	22	US-10-363-483A-12022	Sequence 12022, A
1098	52.4	3.1	679	15	US-10-198-846-2750	Sequence 2750, Ap
1099	52.4	3.1	2309	21	US-10-425-115-47194	Sequence 47194, A
c1100	52.4	3.1	5586	18	US-10-257-166-86	Sequence 86, Appl
1101	52.2	3.1	571	20	US-10-021-323-7115	Sequence 7115, Ap
1102	52.2	3.1	1956	21	US-10-424-599-65202	Sequence 65202, A
1103	52.2	3.1	2010	26	US-11-097-143-14513	Sequence 14513, A
1104	52.2	3.1	2057	18	US-10-266-829-15	Sequence 15, Appl
1105	52.2	3.1	2057	19	US-10-878-523-15	Sequence 15, Appl
1106	52.2	3.1	3063	20	US-10-467-595-83	Sequence 83, Appl
1107	52.2	3.1	3927	14	US-10-108-605-48	Sequence 48, Appl
c1108	52.2	3.1	18218	16	US-10-311-455-1921	Sequence 1921, Ap
1109	52	3.1	183	16	US-10-102-524-1621	Sequence 1621, Ap
1110	52	3.1	484	21	US-10-357-930-58727	Sequence 58727, A
c1111	52	3.1	517	21	US-10-425-115-82733	Sequence 82733, A
1112	52	3.1	648	21	US-10-425-115-140307	Sequence 140307,
c1113	51.8	3.1	5739	16	US-10-311-455-692	Sequence 692, App
1114	51.8	3.1	633	21	US-10-425-115-8940	Sequence 8940, Ap
1115	51.8	3.1	668	19	US-10-424-599-61809	Sequence 61809, A
1116	51.6	3.1	731	21	US-10-425-115-18416	Sequence 18416, A
1118	51.6	3.1	1167	21	US-10-425-115-88123	Sequence 88123, A
1119	51.6	3.1	461	20	US-10-437-963-92730	Sequence 92730, A
1120	51.6	3.1	2911	18	US-10-264-049-283	Sequence 283, Appl
c1121	51.6	3.0	3508	19	US-09-616-263-3	Sequence 3, Appli
1122	51.6	3.1	8759	21	US-10-473-126-249	Sequence 249, App
c1123	51.6	3.1	8759	21	US-10-473-126-377	Sequence 377, App
1124	51.6	3.1	8979	16	US-10-311-455-757	Sequence 757, App
1125	51.4	3.1	8979	18	US-10-221-613-137	Sequence 137, App
1126	51.4	3.0	421	21	US-10-021-323-7658	Sequence 7658, Ap
1127	51.4	3.0	528	20	US-10-425-115-102437	Sequence 102437,
1128	51.4	3.0	585	20	US-10-021-323-4901	Sequence 4901, Ap
c1129	51.2	3.0	324	19	US-10-425-115-149581	Sequence 149581,
1130	51.2	3.0	343	18	US-10-424-599-89995	Sequence 89995, A
1131	51.2	3.0	392	21	US-10-357-930-48375	Sequence 48375, A
1132	51.2	3.0	443	21	US-10-425-115-20056	Sequence 20056, A
1133	51.2	3.0	491	21	US-10-425-115-38412	Sequence 38412, A
1134	51.2	3.0	504	21	US-10-357-930-58645	Sequence 58645, A
1135	51.2	3.0	608	19	US-10-424-599-59595	Sequence 59595, A
1136	51.2	3.0	807	11	US-10-425-115-119727	Sequence 119727,
1137	51.2	3.0	894	20	US-09-764-875-44	Sequence 44, Appl
1138	51.2	3.0	972	21	US-10-425-115-51692	Sequence 51692, A
c1139	51	3.0	1110	20	US-10-437-963-46659	Sequence 46659, A
1140	51	3.0	1381	22	US-10-767-701-9537	Sequence 9537, Ap
1141	51	3.0	1981	21	US-10-425-115-178086	Sequence 178086,
1142	51	3.0	2286	21	US-10-887-553A-930	Sequence 930, App
1143	50.8	3.0	498	19	US-10-425-115-27352	Sequence 27352, A
c1144	50.8	3.0	517	20	US-10-021-323-11054	Sequence 11054, A
1145	50.8	3.0	1495	21	US-10-424-599-36589	Sequence 36589, A
1146	50.8	3.0	409	21	US-10-437-963-93518	Sequence 93518, A
1147	50.8	3.0	540	21	US-10-425-115-117119	Sequence 117119,
1148	50.8	3.0	553	21	US-10-425-115-112028	Sequence 112028,
1149	50.8	3.0	597	15	US-10-425-115-35417	Sequence 35417, A
1150	50.8	3.0	643	21	US-10-198-846-8987	Sequence 8987, Ap
1151	50.8	3.0	845	21	US-10-425-115-60939	Sequence 60939, A
1152	50.8	3.0	1004	21	US-10-425-115-75154	Sequence 75154, A
1153	50.8	3.0	1026	18	US-10-242-535A-53529	Sequence 53529, A
1154	50.8	3.0	1026	19	US-10-085-783A-53529	Sequence 53529, A

#	Score		Len	Accession	Len	Description
1155	50.8	3.0	1661	US-10-425-115-151054	1661	Sequence 151054, A
1156	50.8	3.0	1985	US-10-424-599-67834	1985	Sequence 67834, A
1157	50.6	3.0	2646	US-10-425-115-31139	2646	Sequence 31139, A
1158	50.6	3.0	336	US-10-424-599-100526	336	Sequence 100526, A
c1159	50.6	3.0	393	US-09-814-353-15254	393	Sequence 15254, A
1160	50.6	3.0	395	US-10-425-115-135930	395	Sequence 135930, Ap
1161	50.6	3.0	496	US-10-424-599-74788	496	Sequence 74788, A
1162	50.6	3.0	538	US-10-424-599-51250	538	Sequence 51250, A
1163	50.6	3.0	600	US-10-972-079-78376	600	Sequence 78376, A
1164	50.6	3.0	856	US-09-770-445-613	856	Sequence 613, App
1165	50.6	3.0	1226	US-10-424-599-43121	1226	Sequence 43121, A
c1166	50.6	3.0	2270	US-10-424-599-44604	2270	Sequence 44604, A
1167	50.6	3.0	2924	US-10-205-219-103	2924	Sequence 103, App
c1168	50.6	3.0	4990	US-10-357-930-24994	4990	Sequence 24994, A
1169	50.6	3.0	4990	US-10-357-930-25389	4990	Sequence 25389, A
1170	50.6	3.0	5001	US-10-198-846-11011	5001	Sequence 11011, A
c1171	50.6	3.0	6029	US-10-240-589C-66	6029	Sequence 66, Appl
c1172	50.6	3.0	3673778	US-10-312-841-1	3673778	Sequence 1, Appli
c1173	50.4	3.0	272	US-09-960-352-6986	272	Sequence 6986, Ap
1174	50.4	3.0	319	US-10-242-535A-15028	319	Sequence 15028, A
1175	50.4	3.0	319	US-10-085-783A-15028	319	Sequence 15028, A
1176	50.4	3.0	590	US-10-425-115-102850	590	Sequence 102850, A
1177	50.4	3.0	1340	US-10-425-115-9974	1340	Sequence 9974, Ap
c1178	50.4	3.0	3687	US-10-602-494-233	3687	Sequence 233, App
1179	50.4	3.0	6668	US-10-311-455-1669	6668	Sequence 1669, Ap
1180	50.2	3.0	201	US-10-357-930-19029	201	Sequence 19029, A
1181	50.2	3.0	383	US-10-357-930-34078	383	Sequence 34078, A
1182	50.2	3.0	383	US-10-357-930-42946	383	Sequence 42946, A
1183	50.2	3.0	384	US-09-814-353-2534	384	Sequence 2534, A
1184	50.2	3.0	384	US-10-357-930-8870	384	Sequence 8870, Ap
1185	50.2	3.0	384	US-10-357-930-18971	384	Sequence 18971, A
1186	50.2	3.0	418	US-10-437-963-79114	418	Sequence 79114, A
1187	50.2	3.0	433	US-10-425-115-19917	433	Sequence 19917, A
1188	50.2	3.0	448	US-10-357-930-48836	448	Sequence 48836, A
1189	50.2	3.0	462	US-10-425-115-111572	462	Sequence 111572, A
1190	50.2	3.0	473	US-10-425-115-115810	473	Sequence 115810, A
1191	50.2	3.0	485	US-10-425-115-144725	485	Sequence 144725, A
1192	50.2	3.0	567	US-10-425-115-88563	567	Sequence 88563, A
1193	50.2	3.0	574	US-10-425-115-179039	574	Sequence 179039, A
c1194	50.2	3.0	575	US-10-424-599-59730	575	Sequence 59730, A
1195	50.2	3.0	577	US-10-425-115-98308	577	Sequence 98308, A
1196	50.2	3.0	629	US-10-437-963-61015	629	Sequence 61015, A
1197	50.2	3.0	732	US-10-437-963-101250	732	Sequence 101250, A
1198	50.2	3.0	821	US-10-425-115-121786	821	Sequence 121786, A
1199	50.2	3.0	1460	US-10-424-599-142655	1460	Sequence 142655, A
c1200	50.2	3.0	6476	US-10-311-455-513	6476	Sequence 513, App
1201	50	3.0	306	US-10-425-115-145784	306	Sequence 145784, A
1202	50	3.0	325	US-10-357-930-59036	325	Sequence 59036, A
1203	50	3.0	458	US-10-425-115-111238	458	Sequence 111238, A
1204	50	3.0	464	US-10-357-930-50950	464	Sequence 50950, A
1205	50	3.0	471	US-10-357-930-54342	471	Sequence 54342, A
1206	50	3.0	520	US-09-918-995-32839	520	Sequence 32839, A
1207	50	3.0	539	US-10-425-115-69113	539	Sequence 69113, A
1208	50	3.0	870	US-09-948-820-12	870	Sequence 12, Appl
1209	50	3.0	870	US-10-613-076-12	870	Sequence 12, Appl
1210	50	3.0	1302	US-10-739-930-5166	1302	Sequence 5166, Ap
1211	50	3.0	1492	US-10-425-115-103271	1492	Sequence 103271,
1212	50	3.0	1509	US-10-425-115-156329	1509	Sequence 156329,
1213	50	3.0	1615	US-10-425-115-149513	1615	Sequence 149513,
1214	50	3.0	1878	US-10-425-115-85400	1878	Sequence 85400, A
1215	50	3.0	3469	US-10-723-860-7734	3469	Sequence 7734, Ap
1216	50	3.0	21537	US-10-311-455-1972	21537	Sequence 1972, Ap
1217	50	3.0	263	US-10-242-535A-8184	263	Sequence 8184, Ap
1218	49.8	3.0	263	US-10-085-783A-8184	263	Sequence 8184, Ap
1219	49.8	3.0	270	US-10-085-783A-707	270	Sequence 707, App
1220	49.8	3.0	270	US-10-242-535A-707	270	Sequence 707, APP
1221	49.8	3.0	311	US-10-357-930-49258	311	Sequence 49258, A
1222	49.8	3.0	367	US-10-425-115-108400	367	Sequence 108400, A
1223	49.8	3.0	376	US-10-357-930-57923	376	Sequence 57923, A
1224	49.8	3.0	380	US-10-437-963-20205	380	Sequence 20205, A
1225	49.8	3.0	381	US-10-357-930-54485	381	Sequence 54485, A
c1226	49.8	3.0	414	US-10-425-115-169784	414	Sequence 169784, A
1227	49.8	3.0	465	US-09-918-995-26264	465	Sequence 26264, A
1228	49.8	3.0	502	US-10-425-115-27329	502	Sequence 27329, A
1229	49.8	3.0	506	US-09-918-995-10105	506	Sequence 10105, A
1230	49.8	3.0	534	US-10-425-115-10497	534	Sequence 10497, A
1231	49.8	3.0	543	US-10-357-930-56504	543	Sequence 56504, A
1232	49.8	3.0	571	US-10-021-323-8505	571	Sequence 8505, Ap
1233	49.8	3.0	600	US-10-956-157-5858	600	Sequence 5858, Ap
1234	49.8	3.0	635	US-10-425-115-11555	635	Sequence 11555, A
c1235	49.8	3.0	689	US-10-425-115-118276	689	Sequence 118276,
1236	49.8	3.0	1048	US-10-351-334-38	1048	Sequence 38, Appl
1237	49.8	3.0	1833	US-10-231-956A-76	1833	Sequence 76, Appl
1238	49.8	3.0	1833	US-10-684-422-254	1833	Sequence 254, App
1239	49.8	3.0	1833	US-10-956-157-623	1833	Sequence 623, App
1240	49.8	3.0	1837	US-10-956-157-75	1837	Sequence 75, Appl
1241	49.8	3.0	2044	US-10-439-703-315	2044	Sequence 315, App
1242	49.8	3.0	2044	US-09-925-302-315	2044	Sequence 315, App
1243	49.8	3.0	2045	US-09-925-302-315	2045	Sequence 2, Appl
1244	49.8	3.0	2045	US-10-013-477-2	2045	Sequence 349, App
1245	49.8	3.0	2496	US-10-106-698-349	2496	Sequence 101, App
1246	49.8	3.0	2496	US-09-745-763-101	2496	Sequence 3, Appl
1247	49.8	3.0	3436	US-10-463-260-3	3436	Sequence 12, Appl
1248	49.8	3.0	3436	US-09-789-561-12	3436	Sequence 3, Appl
1249	49.8	3.0	4080	US-09-790-622-3	4080	Sequence 3, Appl
1250	49.8	3.0	3436	US-10-141-953-3	3436	Sequence 12, Appl
1251	49.8	3.0	3436	US-10-883-936-12	3436	Sequence 115754
1252	49.8	3.0	4080	US-10-424-599-115754	4080	Sequence 318, App
c1253	49.6	3.0	7346	US-10-311-455-318	7346	Sequence 1102, Ap
1254	49.6	3.0	353	US-10-437-963-1102	353	Sequence 44608, A
1255	49.6	3.0	378	US-10-425-115-44608	378	Sequence 3585, Ap
c1256	49.6	3.0	414	US-10-425-115-3585	414	Sequence 14917, A
1257	49.6	3.0	481	US-10-021-323-14917	481	Sequence 117433,
c1258	49.6	3.0	543	US-10-425-115-117433	543	Sequence 82545, A
1259	49.6	3.0	613	US-10-425-115-82545	613	Sequence 24054, A
1260	49.6	3.0	660	US-10-424-599-24054	660	Sequence 180319,
1261	49.6	3.0	665	US-10-425-115-180319	665	Sequence 115631,
1262	49.6	3.0	687	US-10-424-599-115631	687	Sequence 422, App
1263	49.6	3.0	716	US-10-764-420-422	716	Sequence 176985,
1264	49.6	3.0	1466	US-10-425-115-176985	1466	Sequence 5536, Ap
1265	49.6	3.0	1486	US-10-723-860-5536	1486	Sequence 128137,
1266	49.6	3.0	1590	US-10-425-115-128137	1590	Sequence 56660, A
1267	49.6	3.0	1743	US-10-424-599-56660	1743	Sequence 98224, A
1268	49.6	3.0	1977	US-10-437-963-98224	1977	Sequence 21598, A
c1269	49.6	3.0	2161	US-09-814-353-21598	2161	Sequence 1491, Ap
1270	49.6	3.0	5520	US-10-311-455-1491	5520	Sequence 650, App
1271	49.6	3.0	6015	US-10-311-455-650	6015	Sequence 2, Appl
c1272	49.6	3.0	3673778	US-10-312-841-2	3673778	Sequence 56685, A
c1273	49.4	2.9	408	US-10-357-930-56685	408	Sequence 45909, A
1274	49.4	2.9	438	US-10-357-930-45909	438	Sequence 42927, A
1275	49.4	2.9	464	US-10-437-963-42927	464	Sequence 126404,
c1276	49.4	2.9	537	US-10-425-115-126404	537	Sequence 53974, A
1277	49.4	2.9	541	US-10-424-599-53974	541	Sequence 12358, A
1278	49.4	2.9	561	US-09-918-995-12358	561	Sequence 58709, A
1279	49.4	2.9	579	US-10-357-930-58709	579	Sequence 181737,
1280	49.4	2.9	642	US-10-425-115-181737	642	Sequence 177175,
1281	49.4	2.9	693	US-10-425-115-177175	693	Sequence 2095, Ap
1282	49.4	2.9	788	US-10-106-698-2095	788	Sequence 9434, Ap
c1283	49.4	2.9	813	US-10-437-963-9434	813	Sequence 86611, A
1284	49.4	2.9	1068	US-10-437-963-86611	1068	Sequence 35301, A
1285	49.4	2.9	1320	US-10-425-115-35301	1320	Sequence 126404,
1286	49.4	2.9	1476	US-10-424-599-80163	1476	Sequence 80163, A
1287	49.4	2.9	1531	US-10-424-599-24007	1531	Sequence 24007, A
1288	49.4	2.9	1531	US-10-437-963-70263	1531	Sequence 70263, A
1289	49.4	2.9	1603	US-10-357-930-22091	1603	Sequence 22091, A
1290	49.4	2.9	1603	US-10-357-930-23103	1603	Sequence 23103, A
1291	49.4	2.9	1603	US-10-357-930-27953	1603	Sequence 27953, A
1292	49.4	2.9	1603	US-10-357-930-28971	1603	Sequence 28971, A
1293	49.4	2.9	1899	US-10-425-115-72710	1899	Sequence 72710, A
1294	49.4	2.9	2645	US-09-764-870-76	2645	Sequence 76, Appl
1295	49.4	2.9	2645	US-10-125-540-76	2645	Sequence 6047, Ap
c1296	49.4	2.9	3480	US-10-723-860-6047	3480	Sequence 7357, Ap
c1297	49.2	2.9	3480	US-10-723-860-7357	3480	Sequence 88721, Ap
1298	49.2	2.9	240	US-10-425-115-88721	240	Sequence 42318, Ap
1299	49.2	2.9	366	US-10-424-599-42318	366	Sequence 8515, Ap
1300	49.2	2.9	377	US-10-425-115-8515	377	Sequence 8515, Ap

Idx		Score	Len	n	US Identifier	Sequence
1301	2.9	49.2	411	19	US-10-424-599-44054	Sequence 44054, A
1302	2.9	49.2	416	21	US-10-425-115-166795	Sequence 166795,
1303	2.9	49.2	442	21	US-10-357-930-58546	Sequence 58546, A
1304	2.9	49.2	448	10	US-09-918-995-12373	Sequence 12373, A
1305	2.9	49.2	463	21	US-10-425-115-166559	Sequence 166559,
1306	2.9	49.2	468	21	US-10-425-115-66234	Sequence 66234, A
1307	2.9	49.2	472	21	US-10-425-115-115781	Sequence 115781,
1308	2.9	49.2	486	19	US-10-424-599-57535	Sequence 57535, A
1309	2.9	49.2	494	20	US-10-437-963-35520	Sequence 35520, A
1310	2.9	49.2	528	21	US-10-425-115-8964	Sequence 8964, Ap
1311	2.9	49.2	539	21	US-10-425-115-167612	Sequence 167612, A
1312	2.9	49.2	539	21	US-10-357-930-58859	Sequence 58859, A
c1313	2.9	49.2	544	21	US-10-357-930-56643	Sequence 56643, A
1314	2.9	49.2	547	20	US-10-437-963-20806	Sequence 20806, A
1315	2.9	49.2	554	21	US-10-357-930-58639	Sequence 58639, A
1316	2.9	49.2	575	21	US-10-357-930-58433	Sequence 58433, A
1317	2.9	49.2	577	21	US-10-425-115-146284	Sequence 146284,
1318	2.9	49.2	639	21	US-10-425-115-39467	Sequence 39467, A
1319	2.9	49.2	687	21	US-10-425-115-44683	Sequence 44683, A
1320	2.9	49.2	712	16	US-10-106-698-1404	Sequence 1404, Ap
1321	2.9	49.2	717	21	US-10-425-115-120791	Sequence 120791,
1322	2.9	49.2	877	21	US-10-491-472-79	Sequence 79, Appl
1323	2.9	49.2	1022	21	US-10-425-115-46658	Sequence 46658, A
1324	2.9	49.2	1134	21	US-10-425-115-176021	Sequence 176021,
1325	2.9	49.2	1209	21	US-10-425-115-89648	Sequence 89648, A
1326	2.9	49.2	1576	21	US-10-723-860-5808	Sequence 5808, Ap
1327	2.9	49.2	2408	21	US-10-723-860-6245	Sequence 6245, Ap
c1328	2.9	49.2	2501	21	US-10-473-126-364	Sequence 364, App
1329	2.9	49.2	2534	19	US-10-424-599-138931	Sequence 138931,
1330	2.9	49.2	3297	21	US-10-437-963-28886	Sequence 28886, A
c1331	2.9	49.2	7058	18	US-10-257-166-133	Sequence 133, Appl
c1332	2.9	49.2	11394	16	US-10-240-453-96	Sequence 96, Appl
c1333	2.9	49.2	35962	21	US-10-473-126-244	Sequence 244, Appl
c1334	2.9	49.2	2140405	14	US-10-027-632-76212	Sequence 76212, A
c1335	2.9	49.2	2140405	18	US-10-027-632-76212	Sequence 76212, A
c1336	2.9	49.2	256	9	US-09-732-560-89	Sequence 89, Appl
1337	2.9	49	342	20	US-10-021-323-9783	Sequence 9783, Appl
1338	2.9	49	347	20	US-10-357-930-58547	Sequence 58547, A
c1339	2.9	49	350	20	US-10-437-963-9525	Sequence 9525, Ap
1340	2.9	49	369	20	US-10-021-323-4742	Sequence 4742, Ap
1341	2.9	49	372	20	US-10-425-115-101053	Sequence 101053,
1342	2.9	49	403	20	US-10-021-323-8207	Sequence 8207, Ap
1343	2.9	49	416	20	US-10-425-115-127223	Sequence 127223,
1344	2.9	49	426	10	US-09-918-995-10337	Sequence 10337, A
1345	2.9	49	429	21	US-10-357-930-57399	Sequence 57399, A
1346	2.9	49	449	21	US-10-425-115-131667	Sequence 131667,
1347	2.9	49	469	20	US-10-424-599-99954	Sequence 99954, A
c1348	2.9	49	469	20	US-10-021-323-6668	Sequence 6668, Ap
1349	2.9	49	469	21	US-10-021-323-13196	Sequence 13196, A
1350	2.9	49	472	20	US-10-357-930-56657	Sequence 56657, A
1351	2.9	49	472	21	US-10-357-930-57039	Sequence 57039, A
c1352	2.9	49	502	21	US-10-425-115-112380	Sequence 112380,
1353	2.9	49	541	20	US-10-357-930-54565	Sequence 54565, A
1354	2.9	49	547	20	US-10-021-323-8671	Sequence 8671, Ap
c1355	2.9	49	549	21	US-10-425-115-78669	Sequence 78669, A
1356	2.9	49	553	21	US-10-363-345A-38653	Sequence 38653, A
1357	2.9	49	553	21	US-10-363-345A-38654	Sequence 38654, A
c1358	2.9	49	553	22	US-10-363-483A-38653	Sequence 38653, A
1359	2.9	49	553	21	US-10-363-483A-38654	Sequence 38654, A
c1360	2.9	49	599	21	US-10-425-115-47514	Sequence 47514, A
1361	2.9	49	781	19	US-10-424-599-113882	Sequence 113882,
1362	2.9	49	863	21	US-10-425-115-184509	Sequence 184509,
1363	2.9	49	932	21	US-10-425-115-105074	Sequence 105074,
1364	2.9	49	1041	20	US-10-767-795-4377	Sequence 4377, Ap
1365	2.9	49	1133	21	US-10-425-115-59679	Sequence 59679, A
1366	2.9	49	1215	21	US-10-425-115-94900	Sequence 94900, A
1367	2.9	49	1480	19	US-10-424-599-86648	Sequence 86648, A
1368	2.9	49	1583	21	US-10-425-115-57459	Sequence 57459, A
1369	2.9	49	1698	21	US-10-723-860-6988	Sequence 6988, Ap
1370	2.9	49	1882	18	US-10-310-154-338	Sequence 338, App
1371	2.9	49	1882	22	US-10-732-923-332	Sequence 332, App
1372	2.9	49	2024	15	US-10-198-846-9931	Sequence 9931, Ap
1373	2.9	49	2413	21	US-10-425-115-150246	Sequence 150246,

Idx	Score		Num	n	US Identifier	Sequence
1374	49	2.9	2936	21	US-10-723-860-7352	Sequence 7352, Ap
1375	49	2.9	2936	21	US-10-723-860-7519	Sequence 7519, Ap
1376	49	2.9	2947	21	US-10-424-599-49805	Sequence 49805, Ap
c1377	49	2.9	3044	18	US-10-264-049-555	Sequence 555, App
c1378	49	2.9	5845	16	US-10-311-455-1635	Sequence 1635, Ap
1379	49	2.9	5930	10	US-09-814-353-19600	Sequence 19600, Ap
c1380	49	2.9	6012	16	US-10-311-455-2031	Sequence 2031, Ap
c1381	49	2.9	6161	16	US-10-311-455-384	Sequence 384, App
c1382	49	2.9	6211	16	US-10-311-455-779	Sequence 779, App
c1383	49	2.9	6522	16	US-10-311-455-1023	Sequence 1023, Ap
1384	49	2.9	11416	16	US-10-311-455-91	Sequence 91, Appl
1385	49	2.9	11416	16	US-10-021-613-19	Sequence 19, Appl
c1386	49	2.9	40862	16	US-10-311-455-2045	Sequence 2045, Ap
c1387	48.8	2.9	289	21	US-10-425-115-100902	Sequence 100902,
c1388	48.8	2.9	299	10	US-09-814-353-4844	Sequence 4844, Ap
1389	48.8	2.9	299	10	US-09-814-353-11141	Sequence 11141, A
1390	48.8	2.9	307	19	US-10-424-599-133076	Sequence 133076,
1391	48.8	2.9	308	21	US-10-357-930-49445	Sequence 49445, A
1392	48.8	2.9	313	21	US-10-357-930-55308	Sequence 55308, A
1393	48.8	2.9	326	21	US-10-357-930-4386	Sequence 4386, Ap
1394	48.8	2.9	340	21	US-10-425-115-38438	Sequence 38438, A
1395	48.8	2.9	382	20	US-10-437-963-5108	Sequence 5108, Ap
1396	48.8	2.9	389	21	US-10-425-115-111841	Sequence 111841,
1397	48.8	2.9	432	21	US-10-425-115-94211	Sequence 94211, A
1398	48.8	2.9	436	20	US-10-021-323-10931	Sequence 10931, A
1399	48.8	2.9	446	21	US-10-357-930-19597	Sequence 19597, A
c1400	48.8	2.9	463	21	US-10-425-115-117603	Sequence 117603,
1401	48.8	2.9	486	20	US-10-021-323-471	Sequence 471, App
1402	48.8	2.9	489	21	US-10-425-115-114216	Sequence 114216,
c1403	48.8	2.9	489	21	US-10-437-963-23682	Sequence 23682, A
1404	48.8	2.9	489	21	US-10-357-930-29580	Sequence 29580, A
1405	48.8	2.9	510	21	US-10-357-930-58133	Sequence 58133, A
1406	48.8	2.9	528	21	US-10-425-115-45126	Sequence 45126, A
1407	48.8	2.9	546	21	US-10-425-115-166637	Sequence 166637,
1408	48.8	2.9	577	21	US-10-437-963-89366	Sequence 89366, A
1409	48.8	2.9	588	20	US-10-437-963-55397	Sequence 55397, A
1410	48.8	2.9	608	21	US-10-425-115-150409	Sequence 150409,
1411	48.8	2.9	611	21	US-10-425-115-89295	Sequence 89295, A
1412	48.8	2.9	673	16	US-10-106-698-1147	Sequence 1147, Ap
1413	48.8	2.9	756	19	US-10-424-599-124469	Sequence 124469,
1414	48.8	2.9	836	10	US-09-978-418-47	Sequence 47, Appl
1415	48.8	2.9	836	20	US-10-485-231-47	Sequence 47, Appl
1416	48.8	2.9	970	21	US-10-425-115-1530	Sequence 1530, Ap
1417	48.8	2.9	980	19	US-10-723-860-4894	Sequence 4894, Ap
1418	48.8	2.9	1129	21	US-10-425-115-88252	Sequence 88252, A
1419	48.8	2.9	1352	21	US-10-425-115-171401	Sequence 171401,
1420	48.8	2.9	1417	21	US-10-425-115-78849	Sequence 78849, A
1421	48.8	2.9	1437	21	US-10-425-115-33071	Sequence 33071, A
1422	48.8	2.9	1685	21	US-10-739-930-1950	Sequence 1950, Ap
1423	48.8	2.9	1813	21	US-10-425-115-53485	Sequence 53485, A
1424	48.8	2.9	1813	9	US-09-880-578-3	Sequence 3, Appli
1425	48.8	2.9	1880	21	US-10-425-115-44753	Sequence 44753, A
1426	48.8	2.9	2231	21	US-10-425-115-164558	Sequence 164558,
1427	48.8	2.9	2428	19	US-10-424-599-88703	Sequence 88703, A
1428	48.8	2.9	3260	20	US-10-437-963-80144	Sequence 80144, A
1429	48.8	2.9	5378	16	US-10-311-455-1852	Sequence 1852, Ap
c1430	48.6	2.9	83391	10	US-10-433-793-124	Sequence 124, App
c1431	48.6	2.9	291	10	US-09-814-353-5133	Sequence 5133, Ap
1432	48.6	2.9	291	19	US-10-424-599-11424	Sequence 11424, A
1433	48.6	2.9	299	19	US-10-424-599-84774	Sequence 84774, A
1434	48.6	2.9	320	21	US-10-357-930-49088	Sequence 49088, A
1435	48.6	2.9	377	20	US-10-437-963-70436	Sequence 70436, A
1436	48.6	2.9	395	21	US-10-437-963-50506	Sequence 50506, A
1437	48.6	2.9	412	21	US-10-425-115-74493	Sequence 74493, A
1438	48.6	2.9	416	19	US-10-424-599-71220	Sequence 71220, A
1439	48.6	2.9	416	21	US-10-425-115-795	Sequence 795, App
1440	48.6	2.9	423	10	US-09-918-995-7147	Sequence 7147, Ap
1441	48.6	2.9	425	21	US-10-357-930-38081	Sequence 38081, A
1442	48.6	2.9	471	21	US-10-357-930-47774	Sequence 47774, A
1443	48.6	2.9	489	20	US-10-021-323-9618	Sequence 9618, Ap
1444	48.6	2.9	496	21	US-10-425-115-61666	Sequence 61666, A
1445	48.6	2.9	500	21	US-10-425-115-90728	Sequence 90728, A
1446	48.6	2.9	516	21	US-10-425-115-155823	Sequence 155823,

1447	48.6	2.9	519	21	US-10-425-115-42803	Sequence 42803, A
1448	48.6	2.9	533	21	US-10-425-115-57335	Sequence 57335, A
c1449	48.6	2.9	539	21	US-10-425-115-126263	Sequence 126263,
1450	48.6	2.9	554	21	US-10-425-115-23306	Sequence 23306, A
1451	48.6	2.9	580	21	US-10-425-115-19983	Sequence 19983, A
1452	48.6	2.9	595	21	US-10-425-115-76922	Sequence 76922, A
1453	48.6	2.9	611	21	US-10-425-115-44206	Sequence 44206, A
1454	48.6	2.9	617	21	US-10-425-115-102811	Sequence 102811,
c1455	48.6	2.9	637	21	US-10-425-115-66531	Sequence 66531, A
1456	48.6	2.9	645	21	US-10-425-115-116990	Sequence 116990,
1457	48.6	2.9	731	21	US-10-425-115-79030	Sequence 79030, A
1458	48.6	2.9	759	19	US-10-424-599-21799	Sequence 21799, A
1459	48.6	2.9	816	21	US-10-357-930-17956	Sequence 17956, A
1460	48.6	2.9	978	21	US-10-425-115-8752	Sequence 8752, Ap
1461	48.6	2.9	1010	21	US-10-425-115-159322	Sequence 159322,
1462	48.6	2.9	1117	19	US-10-424-599-47804	Sequence 47804, A
1463	48.6	2.9	1204	19	US-10-424-599-96067	Sequence 96067, A
1464	48.6	2.9	1377	9	US-09-925-299-93	Sequence 93, Appl
1465	48.6	2.9	1377	10	US-09-925-299-93	Sequence 93, Appl
1466	48.6	2.9	1542	21	US-10-739-930-2419	Sequence 2419, Ap
1467	48.6	2.9	1808	9	US-09-989-920-46	Sequence 46, Appl
1468	48.6	2.9	1965	21	US-10-425-115-141882	Sequence 141882,
1469	48.6	2.9	2034	17	US-10-047-855-2	Sequence 2, Appli
1470	48.6	2.9	2034	18	US-10-426-776-19	Sequence 19, Appl
1471	48.6	2.9	2050	21	US-10-723-860-6437	Sequence 6437, Ap
1472	48.6	2.9	2390	24	US-10-756-149-1338	Sequence 1338, Ap
c1473	48.6	2.9	2818	21	US-10-602-494-120	Sequence 120, App
1474	48.6	2.9	2818	21	US-10-602-494-126	Sequence 126, App
c1475	48.6	2.9	2818	21	US-10-602-494-262	Sequence 262, App
c1476	48.6	2.9	2818	21	US-10-602-494-268	Sequence 268, App
c1477	48.6	2.9	5532	16	US-10-311-455-751	Sequence 751, App
1478	48.6	2.9	8946	16	US-10-311-455-884	Sequence 884, App
c1479	48.6	2.9	14615	19	US-10-221-714A-429	Sequence 429, App
1480	48.4	2.9	186	18	US-10-242-535A-22399	Sequence 22399, A
1481	48.4	2.9	186	19	US-10-085-783A-22399	Sequence 22399, A
c1482	48.4	2.9	226	21	US-10-425-115-155851	Sequence 155851,
1483	48.4	2.9	232	21	US-10-425-115-144728	Sequence 144728,
1484	48.4	2.9	262	19	US-10-424-599-11870	Sequence 11870, A
1485	48.4	2.9	284	19	US-10-424-599-65419	Sequence 65419, A
c1486	48.4	2.9	341	9	US-09-960-352-12302	Sequence 12302, A
1487	48.4	2.9	341	21	US-10-425-115-14890	Sequence 14890, A
1488	48.4	2.9	352	21	US-10-425-115-108824	Sequence 108824,
1489	48.4	2.9	362	21	US-10-425-115-36607	Sequence 36607, A
1490	48.4	2.9	367	21	US-10-425-115-59904	Sequence 59904, A
1491	48.4	2.9	378	21	US-10-425-115-133892	Sequence 133892,
c1492	48.4	2.9	388	10	US-09-814-353-17525	Sequence 17525, A
1493	48.4	2.9	397	21	US-10-425-115-157161	Sequence 157161,
1494	48.4	2.9	400	21	US-10-425-115-109342	Sequence 109342,
1495	48.4	2.9	401	19	US-10-424-599-139036	Sequence 139036,
1496	48.4	2.9	411	21	US-10-425-115-92368	Sequence 92368, A
1497	48.4	2.9	432	21	US-10-357-930-56856	Sequence 56856, A
c1498	48.4	2.9	442	9	US-09-960-352-12911	Sequence 12911, A
1499	48.4	2.9	451	21	US-10-357-930-56797	Sequence 56797, A
1500	48.4	2.9	456	20	US-10-437-963-21602	Sequence 21602, A

Search completed: October 30, 2005, 07:42:38
Job time : 1530 secs

```
            GenCore version 5.1.6
        Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:     October 30, 2005, 10:07:00 ; Search time 961 Seconds
                                          (without alignments)
                          16003.501 Million cell updates/sec

Title:          US-10-009-002-1
Perfect score:  9399
Sequence:       1 accacaaacactccagtttg.........cccgcctggaattaaaaact 9399

Scoring table:  IDENTITY NUC
                Gapop 10.0 , Gapext 1.0

Searched:       1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters:    2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      Issued Patents NA:*
                1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
                6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Match	Query Match Length	DB	ID	Description
1	9059.8	96.4	9143	2	US-08-639-857-32	Sequence 32, Appl
2	9059.8	96.4	9143	2	US-08-469-260A-390	Sequence 390, App
3	9059.8	96.4	9143	3	US-08-469-260A-393	Sequence 393, App
4	9059.8	96.4	9143	4	US-08-488-446-390	Sequence 390, App
5	9059.8	96.4	9143	4	US-08-488-446-393	Sequence 393, App
6	9059.8	96.4	9143	4	US-08-467-344A-390	Sequence 390, App
7	9059.8	96.4	9143	4	US-08-467-344A-393	Sequence 393, App
8	9059.8	96.4	9143	4	US-08-424-550B-390	Sequence 390, App
9	9059.8	96.4	9143	4	US-08-424-550B-393	Sequence 393, App
10	8692.6	92.5	8912	3	US-08-469-260A-11	Sequence 11, Appl
11	8692.6	92.5	8912	4	US-08-488-446-11	Sequence 11, Appl
12	8692.6	92.5	8912	4	US-08-467-344A-11	Sequence 11, Appl
13	8692.6	92.5	8912	4	US-08-424-550B-11	Sequence 11, Appl
14	4159	44.2	4268	3	US-08-469-260A-80	Sequence 80, Appl
15	4159	44.2	4268	4	US-08-488-446-80	Sequence 80, Appl
16	4159	44.2	4268	4	US-08-467-344A-80	Sequence 80, Appl
17	4159	44.2	4268	4	US-08-424-550B-80	Sequence 80, Appl
18	449.8	4.8	479	3	US-08-469-260A-25	Sequence 25, Appl
19	449.8	4.8	479	4	US-08-469-260A-392	Sequence 392, App
20	449.8	4.8	479	4	US-08-488-446-25	Sequence 25, Appl
21	449.8	4.8	479	4	US-08-488-446-392	Sequence 392, App
22	449.8	4.8	479	4	US-08-467-344A-25	Sequence 25, Appl
23	449.8	4.8	479	4	US-08-467-344A-392	Sequence 392, App
24	449.8	4.8	479	4	US-08-424-550B-25	Sequence 25, Appl
25	449.8	4.8	479	4	US-08-424-550B-392	Sequence 392, App
26	319.6	3.4	337	3	US-08-469-260A-29	Sequence 29, Appl
27	319.6	3.4	337	3	US-08-469-260A-159	Sequence 159, App
c 28	319.6	3.4	337	4	US-08-488-446-29	Sequence 29, Appl
c 29	319.6	3.4	337	4	US-08-488-446-159	Sequence 159, App
c 30	319.6	3.4	337	4	US-08-467-344A-29	Sequence 29, App
c 31	319.6	3.4	337	4	US-08-467-344A-159	Sequence 159, App
c 32	319.6	3.4	337	4	US-08-424-550B-29	Sequence 29, Appl
c 33	319.6	3.4	337	4	US-08-424-550B-159	Sequence 159, App
34	307.4	3.3	309	4	US-09-579-302-1	Sequence 1, Appl
35	279.4	3.0	281	3	US-08-469-260A-18	Sequence 18, Appl
36	279.4	3.0	281	4	US-08-469-260A-20	Sequence 20, Appl
37	279.4	3.0	281	4	US-08-488-446-18	Sequence 18, Appl
38	279.4	3.0	281	4	US-08-488-446-20	Sequence 20, Appl
39	279.4	3.0	281	4	US-08-467-344A-18	Sequence 18, Appl
40	279.4	3.0	281	4	US-08-467-344A-20	Sequence 20, Appl
41	279.4	3.0	281	4	US-08-424-550B-18	Sequence 18, Appl
42	279.4	3.0	281	4	US-08-424-550B-20	Sequence 20, Appl
43	265.6	2.8	3970	1	US-07-925-695-3	Sequence 3, Appli
44	265.6	2.8	9589	1	US-07-925-695-1	Sequence 1, Appli
45	265.6	2.8	9589	1	US-07-925-695-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-639-857-32

; Sequence 32, Application US/08639857
; Patent No. 5955318
; GENERAL INFORMATION:
;  APPLICANT:  Simons, J. N.
;  APPLICANT:  Desai, S. M
;  APPLICANT:  Mushahwar, I. K.
;  TITLE OF INVENTION:  REAGENTS AND METHODS USEFUL FOR CONTROLLING THE
;  TITLE OF INVENTION:  TRANSLATION OF HEPATITIS GB PROTEINS
;  NUMBER OF SEQUENCES: 32
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  Abbott Laboratories
;   STREET:  100 Abbott Park Rd
;   CITY:  Abbott Park
;   STATE:  IL
;   COUNTRY:  USA
;   ZIP:  60064
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Floppy disk
;   COMPUTER:  IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/639,857
;   FILING DATE:
;   CLASSIFICATION:  435
;  ATTORNEY/AGENT INFORMATION:
;   NAME:  Porembski, Priscilla E.
;   REGISTRATION NUMBER:  33,207
;   REFERENCE/DOCKET NUMBER:  5793.US.P1
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  708-937-0378
;   TELEFAX:  708-938-2623
; INFORMATION FOR SEQ ID NO:  32:
;  SEQUENCE CHARACTERISTICS:
;   LENGTH:  9143 base pairs
;   TYPE:  nucleic acid
;   STRANDEDNESS:  double
;   TOPOLOGY:  linear
;  MOLECULE TYPE:  DNA (genomic)
US-08-639-857-32

Query Match             96.4%;  Score 9059.8;  DB 2;  Length 9143;
Best Local Similarity   99.6%;  Pred. No. 0;
Matches 9096;  Conservative 0;  Mismatches 37;  Indels 4;  Gaps 1;

Qy      1 ACCACAAACACTCCAGTTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG 60
Db      1 ACCACAAACACTCCAGTTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG 60
```

```
Qy    61  CAGGGGCGTGGGGGATTTCCCCTGCCCCGTCTGCGAGAAGGGTGGAGCCAACCACCTTAGTAT  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  CAGGGGCGTGGGGGATTTCCCCTGCCCCGTCTGCGAGAAGGGTGGAGCCAACCACCTTAGTAT  120

Qy   121  GTAGGCGCGGGGACTCATGAGACGCTCGGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GTAGGCGCGGGGACTCATGAGACGCTCGGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180

Qy   181  CCTGATGGGCGTTCATGGGTTCGTGGTGGTGGTGCGGCCTTTAGGCAGCCGTTCCACGCCCACCA  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  CCTGATGGGCGTTCATGGGTTCGTGGTGGTGGTGCGGCCTTTAGGCAGCCGTTCCACGCCCCACCA  240

Qy   241  CCTCCCAGATAGAGCGGCGGCGGCACTGTAGGGAGGAAGACCGGGACCGGGTCACTACCAAGGACG  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  CCTCCCAGATAGAGCGGCGGCGGCACTGTAGGGAGGAAGACCGGGACCGGGTCACTACCAAGGACG  300

Qy   301  CAGACCTCTTTTGAGTATCAGCGCCTCCGGAGTAGTTGGGCAAGCCCACCTATATGTGT  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  CAGACCTCTTTTGAGTATCAGCGCCTCCGGAGTAGTTGGGCAAGCCCACCTATATGTGT  360

Qy   361  TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCTGATACGGGTCCTTGCCGAGGGGAT  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCCGAGGGGAT  420

Qy   421  CTGGGGAGTCTCGTAGACCGTAGACACATGCCTGTTATTCTACTCAAACAAGTCCTGTACC  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  CTGGGGAGTCTCGTAGACCGTAGACACATGCCTGTTATTCTACTCAAACAAGTCCTGTACC  480

Qy   481  TGCGCCCAGAACGCGCAAGAACAAGCAGCAGGCTTCATATCCTGTGTCATTAAAAC  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  TGCGCCCAGAACGCGCAAGAACAAGCAGCAGGCTTCATATCCTGTGTCATTAAAAC  540

Qy   541  ATCTGTTGAAAGGGGACAACAGCGAGCCAAAGCGCAAAGTCCAGCGCGATGCTCGGCGTTGCCAGC  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  ATCTGTTGAAAGGGGACAACAGCGAGCCAAAGCGCAAAGTCCAGCGCGATGCTCGGCGTTGCCAGC  600

Qy   601  TTTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  TTTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC  660

Qy   661  TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTCGCAATCTTGGAATCCTTCTGGA  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTCGCAATCTTGGAATCCTTCTGGA  720

Qy   721  TTACCCTTTGGGGTGATTGGTGATGTTACAACTCACACAGCCTCTAGTAGGCCCGCTGT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  TTACCCTTTGGGGTGATTGGTGATGTTACAACTCACACAGCCTCTAGTAGGCCCGCTGT  780

Qy   781  GGCAGGAGCGCGGTCGTTGACCAGTCTGCCAGTAGTACGCTTGCTGCGAGGATGGAGTCAA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  GGCAGGAGCGCGGTCGTTGACCAGTCTGCCAGTAGTACGCTTGCTGCGAGGATGGAGTCAA  840

Qy   841  CTGGGGCTACTGGTTGGTTCGGTGTCCCACCTTTGTGTGGTATGTCGCTGCTATCTTTGGCCTG  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  CTGGGGCTACTGGTTGGTTCGGTGTCCCACCTTTGTGTGGTATGTCGCTGCTATCTTTGGCCTG  900

Qy   901  TCCCTGTAGTGGTGGGGCGCGGGGTCACTGACCAGACACAAATACCACCAAATTG  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  TCCCCTGTAGTGGTGGGGCGCGGACCGAGAGTGCGTGGGGTCACTGACCAGACACAAATACCACCAAATTG  960

Qy   961  CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCTTCCCACTTGCCTACACGGCTGGGTTG  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCTTCCCACTTGCCTACACGGCTGGGTTG  1020

Qy  1021  TGTGATCTGCGCGGACGAGTGCTGCGGGTGGGGTTCCCGCCAATCCGTACATCTTCACACCCTTCCAA  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  TGTGATCTGCGCGGACGAGTGCTGCGGGTGGGGTTCCCGCCAATCCGTACATCTTCACACCCTTCAA  1080

Qy  1081  TTGGACTTGGCACCGGACTCCTTCTTCTTGGCTGACCACATTGATTTGTTATGGGGCTCTTGT  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  TTGGACTTGGCACCGGACTCCTTCTTCTTGGCTGACCACATTGATTTGTTATGGGGCTCTTGT  1140


Qy  1141  GACCTGTGACGCCCTTGACATTGGTGAGTTGTGTGCGTGTATTAGTCGGTGACTG  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  GACCTGTGACGCCCTTGACATTGGTGAGTTGTGTGCGTGTATTAGTCGGTGACTG  1200

Qy  1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGA  1260

Qy  1261  AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT  1320

Qy  1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACCAAGTACCATACGCTATTGCGACTAT  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACCAAGTACCATACGCTATTGCGACTAT  1380

Qy  1381  GTTTAGCAGTGTACACTACCTGGCGTCTCGATCTACTACTATGCCTCTCGGGGCAA  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  GTTTAGCAGTGTACACTACCTGGCGTCTCGATCTACTACTATGCCTCTCGGGGCAA  1440

Qy  1441  GTGGTATCAGTTGCTCCTAGCGCTTATGCTTTACATAGAAGCGACCTCTGGAAACCCCAT  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  GTGGTATCAGTTGCTCCTAGCGCTTATGCTTTACATAGAAGCGACCTCTGGAAACCCTAT  1500

Qy  1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTTGCTCTGCCCTTTGATGATACCATGTCC  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTTGCTCGCCCTTTGATGATACCATGTCC  1560

Qy  1561  TTGCCACTCTTATTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC  1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  TTGCCACTCTTATTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC  1620

Qy  1621  CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG  1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  CAGGCCTGTCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG  1680

Qy  1681  TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCTATTCGAA  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCTATTCGCAA  1740

Qy  1741  TGTGCCATCGTACTGCCACTCATGGCACTGATGCAGTGTTGGAACGACCACTCGCAACACTTA  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  TGTGCCATCGTACTGCCACTCATGGCACTGATGCAGTGTTGGAACGACCACTCGCAACACTTA  1800

Qy  1801  CGAAGCATGCGGTGTAAACACCATGGCTAACACCGCTAACGGGCTCAGCCCTGAA  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  CGAAGCATGCGGTGTAAACACCATGGCTAACACCGCTAACGGGCTCAGCCCTGAA  1860

Qy  1861  ATTGGCTATATTACAATACCCGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC  1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  ATTGGCTATATTACAATACCCGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC  1920

Qy  1921  AGGCCATTTGTATTTTGAGGGATCAGATACCCCTATAGTTTATTACTTTTATGACCCTGTGAA  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  AGGCCATTTGTATTTTGAGGGATCAGATACCCCTATAGTTTATTACTTTTATGACCCTGTGAA  1980

Qy  1981  TTCCACTCCTACCACCGGAGAGGTGGGCTAGTTGCCCGGTACCCCACCGTGTGTACG  2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  TTCCACTCCTACCACCGGAGAGGTGGGCTAGTTGCCCGGTACCCCACCGTGTGTACG  2040

Qy  2041  TGGTTCTTGGTTACAGGTTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTCACACAGG  2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  TGGTTCTTGGTTACAGGTTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTCACACAGG  2100

Qy  2101  ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC  2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC  2160

Qy  2161  TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGCTGTCTCAATTCTGTTGTGGGGTTGTGTGG  2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTGGCTGTCTCAATTCTGTTGTGGGGTTGTGTGG  2220

Qy  2221  CAGCAAGTATCTTATTTTAGCCTACCTCTGTCTACTTGTCCCTTTGTTTGGGCGCGGCTTC  2280
```

Left column:

```
Db        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2221  CAGCAAGTATCTTATTTGGCTCGTCCTGTGCTCCCATCCTGTTACTTGTCGTGTACTGTTC  2280

Db  2281  TGGTTACCCTTTGGCGTCCTGTGCTCCCATCCCAGTCGTATCTCCAGTTCCAAGCTGGGATGT  2340
Qy  2281  TGGTTACCCTTTGGCGTCCTGTGCTCCCATCCCAGTCGTATCTCCAGTTCCAAGCTGGGATGT  2340

Db  2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTGCTTTGATTTCTTCATCTGTTGCTATCTCCG  2400
Qy  2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTGCTTTGATTTCTTCATCTGTTGCTATCTCCG  2400

Db  2401  CTGCAGGCTAGTTATGTCTGCCCTTTAGGGTTGTGTGCCCATGGCTTGCCCCT  2460
Qy  2401  CTGCAGGCTACGTTATGTCTGCCCTTTAGGGTTTGTGCCCCATGGCTTGCCCCT  2460

Db  2461  AACTTTCTTTGTTGCAGCAGCTGCTGCCCACCAGATTATGACTGGTGGGTGCGACTGCT  2520
Qy  2461  AACTTTCTTTGTTGCAGCAGCTGCTGCCCACCAGATTATGACTGGTGGGTGCCGACTGCT  2520

Db  2521  AGTGGCAGGGTTAGTTGTTGTGGGCCGGCCTGTAACCGTGGTCACCGTAGCTGCTTGT  2580
Qy  2521  AGTGGCAGGGTTAGTTGTTGTGGGCCGGCCTGTGACCGTGGTCCACGTAGCTGCTTGT  2580

Db  2581  AGGTCCTTGGCCTCTGGTAGCCGTTTAACCCTCTTTGGCTTGGTTACGCCGTCTCAGC  2640
Qy  2581  AGGTCCTTGGCCTCTGGTAGCCGTTTAACCCTCTTTGGCTTTGGTTACGCCGTCTCAGC  2640

Db  2641  TTTTGATACCGGAGATAATTGGAAGGGCTGACACAATACCACCTGTAGTAGCATTGTCAT  2700
Qy  2641  TTTTGACACCGGAGATAATTGGAAGGGCTGACACAATACCACCCGTAGTAGCATTGTCAT  2700

Db  2701  GTCTCGTTTTGGCTTCACTTGTTGCTGCTGTTACCTGCGTGTTAGTTAACTCCTATCT  2760
Qy  2701  GTCTCGTTTTGGCTTCACTTGTTGCTGCTGTTACCTGCGTGTTAGTTAACTCCTATCT  2760

Db  2761  TTGGCAACGTTGGGAGAATTGGTTTGGAACGTTACACTAAGACCGGAGAGTTTTCCT  2820
Qy  2761  TTGGCAACGTTGGGAGAATTGGTTTGGAACGTTACACTAAGACCGGAGAGTTTCTCCT  2820

Db  2821  TGTGCTCGGTTGTTTCCCCGGTGGACAATGACGCGCTGACTTTCTGTGACTTCTGTGTGTCA  2880
Qy  2821  TGTGCTCGGTTGTTTCCCCGGTGGACAATGACACGCTGACTTTCTGTGACTTCTGTGTGTCA  2880

Db  2881  CGTAGCTCTTCTATGTTAACATCCAGTGCAGCATCCGTCTTCTTTGGGACTGACTCTAGGGT  2940
Qy  2881  CGTAGCTCTTCTATGTTAACATCCAGTGCAGCATCCGTCTTCTTTGGGACTGACTCTAGGGT  2940

Db  2941  TAGGGCCCATAGAATGTTGGTGCGTCTCGGAAAGTGTCATGCTCGGTATTCTCATTATGT  3000
Qy  2941  TAGGGCCCATAGAATGTTGGTGCGTCTCGGAAAGTGTCATGCTCGGTATTCTCATTATGT  3000

Db  3001  TCTTAAGTGTTTTCCTCTAGTGTTTGGTGTGAAATGTGTGTTTTCTATAAGCAGCTTGCA  3060
Qy  3001  TCTTAAGTGTTTTCCTCTAGTGTTTGGTGTGAAATGTGTGTTTTCTATAAGCAGCTTGCA  3060

Db  3061  GGTTGATGGTCTTGCCTAATGGGTCTCCGGCGCCCTTGTTCGCGGTCCAGGGTTGGCTAT  3120
Qy  3061  GGTTGATGGTCTTGCCTAATGGGTCTCCGGCGCCCTTGTTCGCGGTCCAGGGTTAGCTAT  3120

Db  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGGAATGAAGGAAGACGCTTGGCGTGTGGGACAC  3180
Qy  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGGAATGAAGGAAGACGCTTGGCGTGTGGGACAC  3180

Db  3181  GCCGCCAGATGGTGGGGCCATTACCCGCCACCTTTACCGCTGACTGTTCTCTTGACCTGG  3240
Qy  3181  GGTTGAGGCAAGGCAAGGGTTGGCGTGTTGTTCGCGGTCCAGGGGTTAGCTAT  3240

Db  3241  GCCGCCAGATGGTGGGGCCATTACCCGCCACCTTTACCGCTGACTGTTCTCTTGACCTGG  3300
Qy  3241  GCCGCCAGATGGTGGGGCCATTACCCGCCACCTTTACCGCTGACTGTCTCTTGACGTGG  3300

Db  3301  CACGCTGTCAGCGATGGCAGTGTCATGACTGGTATAGAGACCCCGAACTTGGACTGGAAC  3360
Qy  3301  CACGCTGTCAGCGATGGCAGTGTCATGACTGGTATAGAGACCCCGAACTTGGACTGGAAC  3360
```

Right column:

```
Db  3301  CACGCTGTCAGCGATGGCAGTGTCATGACTGGTATAGAGACCCCGAACTTGGACTGGAAC  3360
Qy  3361  TATCTTCAGATTAGGATCTCTGGCCCACTAGCTACATGGGATTTGTTGTGACAACGTGTT  3420
Db  3361  TATCTTCAGATTAGGATCTCTGGCCCACTAGCTACACGGGATTTGTTGTGACAACGTGTT  3420

Qy  3421  GTATACTGCTCACCATGGCAGCAAGAGGGCGCCGGTTGGCTCATCCCACAGGCTCTATACA  3480
Db  3421  GTATACTGCTCACCATGGCAGCAGCAAGAGGGCGCCGGTTGGCTCATCCCACAGGCTCATACA  3480

Qy  3481  CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACAACCACCATGTGGAGCTGG  3540
Db  3481  CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACAACCACCATGTGGAGCTGG  3540

Qy  3541  GTCCCTACTCGGTGCTCTTGCGGGGTATCTGGGGGAGACCAAGGGTATCTGGTAACGACTGGGGTC  3600
Db  3541  GTCCCTACTCGGTGCTCTTGCGGGGTATCTGGGGGAGACCAAGGGTATCTGGTAACGACTGGGGTC  3600

Qy  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGGTGCGGGGCCCTTCCCAT  3660
Db  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGGTGCGGGGCCCTTCCCAT  3660

Qy  3661  GGCTGTGTGCCAAGGGTCTTCAGGTGCCCGATTCTGTGCTCCTCGGGCATGTTATTGG  3720
Db  3661  GGCTGTGTGCCAAGGGTCTTCAGGTGCCCGATTCTGTGCTCCTCGGGCATGTTATTGG  3720

Qy  3721  GATGTTCACCGCTCTAGAAATTCTGACGGTTCAGTCAGTTCAGGGTTAGGCCGTT  3780
Db  3721  GATGTTCACCGCTCTAGAAATTCTGACGGTTCAGTCAGTTCAGGGTTAGGCCGTT  3780

Qy  3781  GGTGTGTGCTGATACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC  3840
Db  3781  GGTGTGTGCTGATACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC  3840

Qy  3841  TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACCGGCAGCGGCAAGTCAAC  3900
Db  3841  TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACCGGCAGCGGCAAGTCAAC  3900

Qy  3901  CAAATTACCACTTCTTACATGCAGGAGAAGTATGAGTCTTGGTCCTAAATCCCAGTGT  3960
Db  3901  CAAATTACCACTTCTTACATGCAGGAGAAGTATGAGTCTTGGTCCTAAATCCCAGTGT  3960

Qy  3961  GGCTACAACAGCATCAATGCCAAAGTACATGCCAAAGTACACGCGGCTACGGCGTGAATCCAAATTG  4020
Db  3961  GGCTACAACAGCATCAATGCCAAAGTACATGCCAAAGTACACGCGGCTACGGCGTGAATCCAAATTG  4020

Qy  4021  CTATTTAATGCAAATGTACCAACACAGGGCTTCACTTAGGTACGTACACACAATGGCAT  4080
Db  4021  CTATTTAATGCAAATGTACCAACACAGGGCTTCACTTAGGTACGTACACACAATGGCAT  4080

Qy  4081  GTACCTGACCGGAGCATGTTCCCGGAACTCAATGTAATCATTGTGACGATGCCATGC  4140
Db  4081  GTACCTGACCGGAGCATGTTCCCGGAACTCAATGTAATCATTGTGACGATGCCATGC  4140

Qy  4141  TACCGATGCAACCACCGTGTTGGGCATTGGAAAGGTCCATCCAAGCTCCATCCAAAAA  4200
Db  4141  TACCGATGCAACCACCGTGTTGGGCATTGGAAAGGTCCATCCAAGCTCCATCCAAAAA  4200

Qy  4201  TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCTGGAGTAATCCCTTACCACCACATGC  4260
Db  4201  TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCTGGAGTAATCCCTTACCACCACATGC  4260

Qy  4261  CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAGAT  4320
Db  4261  CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAGAT  4320

Qy  4321  TAAGGAGGAAAATCTGAAGAAGGGAGACACCTTTATCTTTGAGGGGAATAAACACTG  4380
Db  4321  TAAGGAGGAAAATCTGAAGAAGGGAGACACCTTTATCTTTGAGGGGAATAAACACTG  4380

Qy  4381  TGATGAGCTTGCTAACGAGTTAGCTGAACAGTCGTCTCTTACTATAGGGG  4440
Db  4381  TGATGAGCTTGCTAACGAGTTAGCTGAACAGTCGTCTCTTACTATAGGGG  4440
```

```
Qy  4441  ATGTGACATCTCAAAAATCCCTGAGGGCGACTGTGTAGTAGTTGCCACTGATGCCTGTG  4500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441  ATGTGACATCTCAAAAATCCCTGAGGGCGACTGTGTAGTAGTTGCCACTGATGCCTGTG  4500

Qy  4501  TACAGGGTACACGTGGTGACTTGACTTGATTCCGTGTATGACTGCAGCCTCATGTGGACAGGCAC  4560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501  TACAGGGTACACGTGGTGACTTGACTTGATTCCGTGTATGACTGCAGCCTCATGTGAGAGGCAC  4560

Qy  4561  ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTGTTCGTGTGTCGGGGTTTCAGC  4620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561  ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTGTTCGTGTGTCGGGGTCTCAGC  4620

Qy  4621  AATAGTTAAAGGCCAGCGTAGGGGCCCGACACAGGCCGTGGGAGAGCTGGCATATACTACTA  4680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621  AATAGTTAAAGGCCAGCGTAGGGGCCCGACACAGGCCGTGGGAGAGCTGGCATATACTACTA  4680

Qy  4681  TGTAGACGGGAGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT  4740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681  TGTAGACGGGAGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT  4740

Qy  4741  CGACGCAGCCAAGGCATGGTATGTTGTCATCAACAGAGCTCAAACTATTCTGGACAC  4800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741  CGACGCAGCCAAGGCATGGTATGTTGTCATCAACAGAGCTCAAACTATTCTGGACAC  4800

Qy  4801  CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT  4860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801  CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT  4860

Qy  4861  CTTTTCTATGGTCAACCCGAACCTTCATTGTCAATACTGCAAAAAGAACTGCTGACAA  4920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861  CTTTTCTATGGTCAACCCGAACCTTCATTGTCAATACTGCAAAAAGAACTGCTGACAA  4920

Qy  4921  TTATGTTTGTTGACTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCC  4980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921  TTATGTTTGTTGACTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCC  4980

Qy  4981  CAATGACGCACCACGGTGGCAGGGAGCCCGGCTTGGGAAAAACCTTGTGGGGTTCTGTG  5040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981  CAATGACGCACCACGGTGGCAGGGAGCCCGGCTTGGGAAAAACCTTGTGGGGTTCTGTG  5040

Qy  5041  GCGCTTGGACGCGCTGACGCGCTGTCCTGCCCAGAGCCCAGCGAGGTGACCAGATACCA  5100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041  GCGCTTGGACGCGCTGACGCGCTGTCCTGCCCAGAGCCCAGCGAGGTGACCAGATACCA  5100

Qy  5101  AATGCTTCACTGAAGTCATAATCATTCCTTGGACAGCCCGACTCGCTGTTGGCGTTGGAGT  5160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101  AATGCTTCACTGAAGTCATAATCATTCCTTGGACAGCCCGACTCGCTGTTGGCGTTGGAGT  5160

Qy  5161  GGCTATGGCTTATCTAGCCATTGACACTTTGACACTTTGGCGCCACTTGTGTGCGGCGTTGCTGTC  5220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161  GGCTATGGCTTATCTAGCCATTGACACTTTGACACTTTGGCGCCACTTGTGTGCGGCGTTGCTGTC  5220

Qy  5221  TATTACATCAGTCCTACCGGTTGCTACTGTCTATCATAGAGTGTATGAAGAAGAAATCGT  5280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221  TATTACATCAGTCCTACCGGTTGCTACTGTCTATCATAGAGTGTATGAAGAAGAAATCGT  5280

Qy  5281  GGAGGAGTTGCATCATTCATTCCCTTGGAGGCCATGGTTGCTGCAATTGACAAGCTGAA  5340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281  GGAGGAGTTGCATCATTCATTCCCTTGGAGGCCATGGTTGCTGCAATTGCGATAAGCTGAA  5340

Qy  5341  GAGTACAATCACCACCAGTCCTTTTGACACTTCACATTGGAAACCGCCCTTGAAAACTTAACAC  5400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341  GAGTACAATCACCACCAGTCCTTTTGACACTTCACATTGGAAACCGCCCTTGAAAAACTTAACAC  5400

Qy  5401  CTTTCTTGGGCCTCATGCGAGCTACACCCTTGCTTATCATAGTATTGCTGTGTTTAGT  5460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401  CTTTCTTGGGCCTCATGCGAGCTACACCCTTGCTTATCATAGTATTGCTGTGTTGGCTTAGT  5460

Qy  5461  CACTTTACCTGACAAATCCCCTTTGCATCATGGGTGTTTGCTTTGCTTCATTGCCGGTATTACTAC  5520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461  CACTTTACCTGACAAATCCCCTTTGCATCATGGGTGTTTGCTTTGCTTCATTGCCGGTATTACTAC  5520
```

```
Qy  5521  CCCACTACCTCACAGATCAAAATGTTCCTGTCATTATTTGGAGGCGCAATTGCCTCCAA  5580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521  CCCACTACCTCACAGATCAAAATGTTCCTGTCATTATTTGGAGGCGCAATTGCCTCCAA  5580

Qy  5581  GCTTACGACGCTAGAGCGCACTGGCGTTCATGATGCGGCGGGCTCGCGGGACAACAGCTCT  5640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581  GCTTACGACGCTAGAGCGCACTGGCGTTCATGATGCGGCGGGCTCGCGGGACAACAGCTCT  5640

Qy  5641  TGGTACATGGACATCGGTGGTTGCTTGTCTTGACATGGTAGGCGGCTATGCTGCCGCCTC  5700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641  TGGTACATGGACATGGACATCGGTGGTTTGTCTTGACATGGTAGGCGGCTATGCTGCCGCCTC  5700

Qy  5701  ATCCACTGCTTGCTTGCTTGACATTAAATGCTTGATGGTGAGTGAGTGCCCACTATGGATCAGCT  5760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701  ATCCACTGCTTGCTTGCTTGACATTAAATGCTTGATGGTGAGTGAGTGCCCACTATGGATCAGCT  5760

Qy  5761  TGCTGGTTAGTGTACTCCGCGTTCAATCCGGCCCGCAGAGGGCCAGAGTTGTGGGGCGTCTTGTCAGC  5820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761  TGCTGGTTAGTGTACTCCGCGTTCAATCCGGCCCGCAGAGGGCCAGAGTTGTGGGGCGTCTTGTCAGC  5820

Qy  5821  TTGTGCAATGTTGCTTGCTTGACAACAGCAGGGCCAGATCACTCGGCCCAACAGACTTCTTAC  5880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821  TTGTGCAATGTTGCTTGCTTGACAACAGCAGGGCCAGATCACTCGGCCCAACAGACTTCTTAC  5880

Qy  5881  TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG  5940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881  TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG  5940

Qy  5941  CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCCTGGAGTGTCATATCAGCTTGCAT  6000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941  CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCCTGGAGTGTCATATCAGCTTGCAT  6000

Qy  6001  CCGTTGGCTCCACACCCCGACCTCATTGCTGCGGCCTCCATGCGGCCGTCGAACTCATCTTTTC  6060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001  CCGTTGGCTCCACACCCCGACCTCATTGCTGCGGCCTCCATGCGGCCGTCGAACTCATCTTTTC  6060

Qy  6061  TTGGCAGTATGTGCAATTCTTTGTGATTGCTTTAATGTCCTTAAAGCTGGAGTTCA  6120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061  TTGGCAGTATGTGCAATTCTTTGTGATTGCTTTAATGTCCTTAAAGCTGGAGTTCA  6120

Qy  6121  GAGCATGTTAACATTCCTGTGTCCTTTCTACAGCTTGCCAGAGGGGTACAAGGGCCC  6180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121  GAGCATGTTAACATTCCTGTGTCCTTTCTACAGCTTGCCAGAGGGGTACAAGGGCCC  6180

Qy  6181  CTGGATTGGATCAGTAGTGCTGCTCCAAGCACGCGTCCATGCGGTTGCTGAACTCATCTTTTTC  6240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181  CCGGTTGGCTCCTCCAAGCACGCGTCCATGCGGTTGCTGAACTCATCTTTTTC  6240

Qy  6241  TGTTGAGAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241  TGTTGAGAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300

Qy  6301  AGGGGCTGTTCCAGTCAACGCTCAGGCTGTGTGTGGCTAGACCGTTGGCGGCCAACCAACTGATTG  6360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301  AGGGGCTGTTCCAGTCAACGCTCAGGCTGTGTGTGGCTAGACCGTTGGCGGCCAACCAACTGATTG  6360

Qy  6361  GACTAGTCTTGTCGTCAATTATGGCGTCAAATATGAGAAAATGGGAGA  6420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361  GACTAGTCTTGTCGTCAATTATGGCGTCAAATATGAGAAAATGGGAGA  6420

Qy  6421  TCACATTTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTTCACCCAGGTGCCCCCAAC  6480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421  TCACATTTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTTCACCCAGGTGCCCCCAAC  6480

Qy  6481  CTTGAGAGCTCGACGGCGTGACGCTTACAGGTTCAGTGTTATCTAGGTGAGCGCCAA  6540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481  CTTGAGAGCTCGACGGCGTGACGCTTACAGGTTCAGTGTTATCTAGGTGAGCGCCAAC  6540

Qy  6541  AACTCCTTGACGACATCCGTGACGGTCAAGGTAAAACTGTTAA  6600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541  AACTCCTTGACGACATCCGTGACGGTCAAGGTAAAACTGTTTAA  6600

Qy  6601  GCTTCCCTTCCGCCGGTGACCGGTGCCCATGCGCAACCTGGTCAACTTAATTTCGTGA  6660
```

```
Db  7681 AGTTCCAAGGAGGAGGAGTCTTCGTGAAGACCCCCAGAAACCAACAAAGAACCCCCAAG 7740

Qy  7741 GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAAGAGATGTACTACGGTCAGT 7800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7741 GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAAGAGATGTACTACGGTCAGT 7800

Qy  7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGATGCGTTGTGGGGTTGTAGATCCACGTAC 7860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGATGCGTTGTGGGGTTGTGACCCACGTAC 7860

Qy  7861 CCGTGTCAAGGCTCTGTTGTCGATGTGGTCACCCGATGCAGTCGGAGCCACATGCGATAC 7920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7861 CCGTGTCAAGGCTCTGTTGTCGATGTGGTCACCCGATGCAGTCGGAGCCACATGCGATAC 7920

Qy  7921 AGTGTGTTTGACAGTACCATCACACACACCCGAGGATATCATGGTGGAGACAGACATCTC 7980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7921 AGTGTGTTTGACAGTACCATCACACACACCCGAGGATATCATGGTGGAGACAGACATCTC 7980

Qy  7981 AGCAGCTAAACTCAGTGACCAACACCAGGCCATTCACACCATTGGCAGGCAGTTATA 8040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7981 AGCAGCTAAACTCAGTGACCAACACCAGGCCATTCACACCATTGGCAGGCAGTTATA 8040

Qy  8041 CGCTGGAGGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGTGTGAGGTC 8100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8041 CGCTGGAGGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGTGTGAGGTC 8100

Qy  8101 TTTCCGGGGTCTATATCTACCTCAAGTTCAACAGTTTGACCTCGCTGGCTGAAGGTAAATGC 8160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8101 TTTCCGGGGTCTATATCTACCTCAAGTTCAACAGTTTGACCTCGCTGGCTGAAGGTAAATGC 8160

Qy  8161 TGCAGCCGAACAGGCTGGCACCACAAGATTGTGTGCCTCAACCCACCCAAATACAGTTGCAC 8220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161 TGCAGCCGAACAGGCTGGCACCACAAGATTGTGTGCCTCAACCCACCCAAATACAGTTGCAC 8220

Qy  8221 CGTAATTGGAAGAGCCGGAGCAGATGCAGACAAGCAAAGCAACAAGCAATGCGTGTTTGCTAG 8280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8221 CGTAATTGGAAGAGCCGGAGCAGATGCAGACAAGCAAAGCAACAAGCAATGCGTGTTTGCTAG 8280

Qy  8281 CTGGATGAAGGTGATGGGTGCACCACAAGATTGTGTGCCTCAACCCACCCAAATACAGTTTGGA 8340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281 CTGGATGAAGGTGATGGGTGCACCACAAGATTGTGTGCCTCAACCCACCCAAATACAGTTTGGA 8340

Qy  8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAGTGGCAAGCCTTA 8400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAGTGGCAAGCCTTA 8400

Qy  8401 CTACTTCTTACAAGAGATCCTGTATCCCCTTGGGCAGGTGCTCTCGCCGAGGGTCTGGG 8460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8401 CTACTTCTTACAAGAGATCCTGTATCCCCTTGGGCAGGTGCTCTCGCCGAGGGTCTGGG 8460

Qy  8461 ATACAACCCAGTGCTGCGGTGGATTGGGTATCTAATAGACAACAACCTCT 8520
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  8461 ATACAACCCAGTGCTGCGGTGGATTGGGTATCTAATAGACAACAACCTCT 8520

Qy  8521 TAGCCGTGTGTGGCTGTCCATTTCATGGAGCAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8521 TAGCCGTGTGTGGCTGTCCATTTCATGGAGCAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580

Qy  8581 GACTGTGACCTTTGACTGGTATGGAGAAAAAATTATACCGGTGCCTTGTGTAGAAGATCTGCCCAG 8640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8581 GACTGTGACCTTTGACTGGTATGGAGAAAAAATTATACCGGTGCCTTGTGTAGAAGATCTGCCCAG 8640

Qy  8641 CATCATTGCTGTGGTGTGCACGGTATTGGGCTTTTCTCGGTGCGCTCACCAACGCTGA 8700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8641 CATCATTGCTGTGGTGTGCACGGTATTGGGCTTTTCTCGGTGCGCTCACCAACGCTGA 8700

Qy  8701 GATCCCAGAGTTCCCAATCACTAACAGACTGACCAATGCCGCCCCCTGCGAGCCTGGCG 8760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8701 GATCCCAGAGTTCCCAATCACTAACAGACTGACCAATGCCGCCCCCTGCGAGCCTGGCG 8760

Qy  8761 AAAGAAAGCCAGGCGCGGTCCTCGCCAAGAGGCGTGGCGGAGCACGCAAAATT 8820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8761 AAAGAAAGCCAGGCGCGGTCCTCGCCAAGAGGCGTGGCGGAGCACGCAAAATT 8820
```

```
Db  6601 ................GCTTCCCTTCCGCGTTGACGGTCACACCTGGTGTGCGCATGCAACTTAATTGCGTGA 6660

Qy  6661 TGCACTTGAGACAAATGACTTGTAATTCCACAAACACTCCAGTGATGAAGCCGCAGT 6720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661 TGCACTTGAGACAAATGACTTGTAATTCCATAAACAACTCCAGTGATGAAGCCGCAGT 6720

Qy  6721 GTCCGGCTCTTGTTTCAAACAGGAGTTGCGCGGTACAAACCACCAATTGCTTGAGGCAATTTC 6780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721 GTCCGGCTCTTGTTTTCAAACAGGAGTTGCGCGGTACAAACCAATTGCTTGAGGCAATTTC 6780

Qy  6781 AGCTGGCGGTTGACACCACCAAACTGCCAGCCCCTCCATCGAAGAGTAGTGGTAAGAAA 6840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781 AGCTGGCGGTTGACACCACCAAACTGCCAGCCCCTCCATCGAAGAGTAGTGGTAAGAAA 6840

Qy  6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTCGCTTCGCTTCCCTCCGAGATCCGTCCC 6900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTCGCTTCGCTTCCCTCCGAGATCCGTCCC 6900

Qy  6901 AGGAGTGTCATGTGCTGAAAGCCTGCAACGAGTGACCGTTGAGAGGTCCTTCAAACCT 6960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901 AGGAGTGTCATGTGCTGAAAGCCTGCAACGAGTGACCGTTGAGAGGTCCTTCAAACCT 6960

Qy  6961 CCCTTCCTCACCACTGTTCTACAGTTGGCCATGCCATGCCCTGTTGGGAGCGGGTGA 7020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961 CCCTTCCTCACCACTGTTCTACAGTTGGCCATGCCATGCCCTGTTGGGAGCGGGTGA 7020

Qy  7021 GTGTAACCCTTCACTGCAATTGCAATTGGATGTGATGACCGAAACAGGCGGAGGCCGGAAGGA 7080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021 GTGTAACCCTTCACTGCAATTGGATGTGATGACCGAAACAGGCGGAGGCCGGAAGGA 7080

Qy  7081 TTTACCCAGTTACCCTCCAAAAAAGGAAGGTCTTACCGGCCCCAAACGGCTACAAAAGAGAGTTGAGTT 7140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081 TTTACCCAGTTACCCTCCAAAAAAGGAAGGTCTTACCGGCCCCAAACGGCTACAAAAGAGAGTTGAGTT 7140

Qy  7141 GGCTACAACCGCTTCCAGCTTACGTTACTCGGCCCCCTACCCGTAAGATACGGGAAGGA 7200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141 GACTACAACCGCTTCCAGCTTACGTTACTCGGCCCCCTACCCGTAAGATACGGGAAGGA 7200

Qy  7201 TTCCAACTCAGTCAGCCCCGCCCAAACGGCCTACAAAAAAGGAAGTTGGGAAGAAGTGAGTT 7260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201 TTCCAACTCAGTCAGCCCCGCCCAAACGGCCTACAAAAAAGGAAGTTGGGAAGAAGTGAGTT 7260

Qy  7261 TTCGGTGCAGCATGATGAGTCGGGGCACTACACCACCAGAAGCAAGCAACTGCTTCAAAGT 7320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261 TTCGGTGCAGCATGATGAGTCGGGGCACTACACCACCAGAAGCAAGCAACTGCTTCAAAGT 7320

Qy  7321 TCTGTCTGCAACTCGGGCACTGCGGGGATGGCCATCCACTAGTGTTTCCTCAAACAAAGATCATTGGTATGT 7380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7321 TCTGTCTGCAACTCGGGCACTGCGGGGATGGCCATCCACTAGTGTTTCCTCAAACAAAGATCATTGGTATGT 7380

Qy  7381 GACTGAGCCGCGGGATGCGGAGCTTAGAAAACAAAAGTCACTATTAATAGACAACAACCTCT 7440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381 GACTGAGCCGCGGGATGCGGAGCTTAGAAAACAAAAGTCACTATTAATAGACAACAACCTCT 7440

Qy  7441 GTTCCCCCCATCATCACCACAAGCAAGATTGGCTAAGGAAAAAGCTTCAAAGTTGT 7500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441 GTTCCCCCCATCATCACCACAAGCAAGATTGGCTAAGGAAAAAGCTTCAAAGTTGT 7500

Qy  7501 CGGGTGTCATGTGACTATGGGACTATGAAGCAGCTAGACGCTCACACGCCCAAACTCTGCTAAGTC 7560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7501 CGGGTGTCATGTGACTATGGGACTATGAAGCAGCTAGACGCTCACACGCCCAAACTCTGCTAAGTC 7560

Qy  7561 GGACTTGCAGAAGTGTGTCGACGCAGTGTGTCGACCCAGTGAGATTATCGGCAAACTGTGAT 7620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7561 CCACTTGCAGAAGTGTGTCGACGCAGTGTGTCCCAATCACTAACAGTTTCCCAATCGTGAT 7620

Qy  7621 GGACTTGCAGAAGTGTGTCGACGCAGTGTGTCGACCCAGTGAGATTATCGGCAAACTGTGAT 7680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7621 GGACTTGCAGAAGTGTGTCGACGCAGTGTGTCCCAATCACTAACAGTTTCCCAATCGTGAT 7680

Qy  7681 AGTTCCAAGGAGGAGGAGTCTTCGTGAAGACCCCCAGAAACCAACAAAGAACCCCCAAG 7740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7681 AGTTCCAAGGAGGAGGAGTCTTCGTGAAGACCCCCAGAAACCAACAAAGAACCCCCAAG 7740
```

```
;       TYPE: nucleic acid
;       STRANDEDNESS: double
;       TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
US-08-469-260A-390

Query Match             96.4%;  Score 9059.8;  DB 3;  Length 9143;
Best Local Similarity   99.6%;  Pred. No. 0;
Matches 9096; Conservative 0; Mismatches 37; Indels 4; Gaps 1;

Qy     1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCTAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCTAG 60

Qy    61 CAGGGCGTGGGGATTTCCCCTGCCCGTCTGCAGAAGGGTGGAGCCAACCACCTTAGTAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CAGGGCGTGGGGATTTCCCCTGCCCGTCTGCAGAAGGGTGGAGCCAACCACCTTAGTAT 120

Qy   121 GTAGGCGCGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTAGGCGCGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC 180

Qy   181 CCTGATGGCGTTCATGGGTTCGGTGGTGGTGGCGCTTTAGGCAGCCTCCACGCCCACCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCTGATGGCGTTCATGGGTTCGGTGGTGGTGGCGCTTTAGGCAGCCTCCACGCCCACCA 240

Qy   241 CCTCCCAGATAGAGCGCGGCGCCACTGTAGGGAGGAGACCGGGACCGGTCACTACCAAGGACG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCCCAGATAGAGCGCGGCGCCACTGTAGGGAGGAGACCGGGACCGGTCACTACCAAGGACG 300

Qy   301 CAGACCTCTTTGAGTATCACGCTCCCGAAGTAGTTGGGCAAGCCCACCTATATGTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGACCTCTTTGAGTATCACGCTCCCGAAGTAGTTGGGCAAGCCCACCTATATGTGT 360

Qy   361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCGTGATAGGTCCTTGCGAGGGGAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCGTGATAGGTCCTTGCGAGGGGAT 420

Qy   421 CTGGGAGTCTCGTAGACCGTAGCACATGCCGTGTTATTTCTACTCAAACAAGTCCTGTACC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTGGGAGTCTCGTAGACCGTAGCACATGCCGTGTTATTTCTACTCAAACAAGTCCTGTACC 480

Qy   481 TGCCCCAGAACGCGCAAGAACAAGCAGACGCAGCTTCATATCCTGTGTCCATTAAAAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TGCCCCAGAACGCGCAAGAACAAGCAGACGCAGCTTCATATCCTGTGTCCATTAAAAC 540

Qy   541 ATCTGTTGAAAGGGACAACGAGCAAAGCGCAAAGTCCAGCGCGATGCTCGGCCTCGTAA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 ATCTGTTGAAAGGGACAACGAGCAAAGCGCAAAGTCCAGCGCGATGCTCGGCCTCGTAA 600

Qy   601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC 660

Qy   661 TCATGGTTGGGGACGCCAAGACCCTCGCCATAGTCTCGCAATCTTGGAATCCTTCTGGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TCATGGTTGGGGACGCCAAGACCCTCGCCATAGTCTCGCAATCTTGGAATCCTTCTGGA 720

Qy   721 TTACCCTTGGGGTGGATTGGTGATGTTACAACTCACACACCTCCAGTAGGGCCCGCTGGT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TTACCCTTGGGGTGGATTGGTGATGTTACAACTCACACACCTCCAGTAGGGCCCGCTGGT 780

Qy   781 GGCAGGAGCGGCTCGTTCGACCAGTCTCGCCAGATAGTACGGCTTGCTGCTGGAGGATGGAGTCAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GGCAGGAGCGGCTCGTTCGACCAGTCTCGCCAGATAGTACGGCTTGCTGCTGGAGGATGGAGTCAA 840

Qy   841 CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTTGTGTATGTCGCTATCTTTGGCCTG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTTGTGTATGTCGCTATCTTTGGCCTG 900

Qy   901 TCCCTGTAGTGGGGCGCGGGTCACTGACCCAGACACAAATACCACAAATCCTGACCAATTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCCTGTAGTGGGGCGCGGGTCACTGACCCAGACACAAATACCACAAATCCTGACCAATTG 960

Qy  8821 GGCTCGCTTCCTTCCTCTGGCATGCTACATCTAGACCTCTACCAGATTGGATAAGACGAG 8880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8821 GGCTCGCTTCCTTCCTCTGGCATGCTACATCTAGACCTCTACCAGATTGGATAAGACCAG 8880

Qy  8881 CGTGGCTCGGTACACCACTTTCAATTATTGTGTGATGTTTACTCCCCGGAGGGGGATGTGTT 8940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8881 CGTGGCTCGGTACACCACTTTCAATTATTGTGTGATGTTTACTCCCCGGAGGGGGATGTGTT 8940

Qy  8941 TATTACACCACAGAGAAGATTGCAGAAGTTCCTTGTGGAAGTATTGGCGTGTCATTGTTTT 9000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8941 TGTTACACCACAGAGAAGATTGCAGAAGTTTCTGTGAAGTATTGGCGTGTCATTGTTTT 9000

Qy  9001 TGCCCTAGGGCTTCATTGCTGTTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9001 TGCCCTAGGGCTTCATTGCTGTTGGACTAGCCATCAGCTGAACCCCAAATTCAAAATTAA 9060

Qy  9061 CTAACAG---TTTTTTTTTTTTTTTTTTTAGGGCAGCCGCAACAGGGGAGACCCC 9116
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCCGCAACAGGGGAGACCCC 9120

Qy  9117 GGGCTTAACGACCCCGC 9133
          |||||||||||||||||
Db  9121 GGGCTTAACGACCCCGC 9137


RESULT 2
US-08-469-260A-390
; Sequence 390, Application US/08469260A
; Patent No. 6451578
; GENERAL INFORMATION:
;       APPLICANT: JOHN N. SIMONS
;       APPLICANT: TAMI J. PILOT-MATIAS
;       APPLICANT: GEORGE J. DAWSON
;       APPLICANT: GEORGE G. SCHLAUDER
;       APPLICANT: SURESH M. DESAI
;       APPLICANT: THOMAS P. LEARY
;       APPLICANT: ANTHONY SCOTT MUERHOFF
;       APPLICANT: JAMES C. ERKER
;       APPLICANT: SHERI L. BUIJK
;       APPLICANT: ISA K. MUSHAHWAR
;       TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
;       TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
;       NUMBER OF SEQUENCES: 716
;       CORRESPONDENCE ADDRESS:
;           ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
;           STREET: 100 ABBOTT PARK ROAD
;           CITY: ABBOTT PARK
;           STATE: IL
;           COUNTRY: USA
;           ZIP: 60064-3500
;       COMPUTER READABLE FORM:
;           MEDIUM TYPE: Floppy disk
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;           SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/08/469,260A
;           FILING DATE:
;           CLASSIFICATION:
;       PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: US/08/424,550
;           FILING DATE:
;       ATTORNEY/AGENT INFORMATION:
;           NAME: POREMBSKI, PRISCILLA E.
;           REGISTRATION NUMBER: 33,207
;           REFERENCE/DOCKET NUMBER: 5527.PC.01
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: 708-937-6365
;           TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 390:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 9143 base pairs
```

The page consists of a nucleotide sequence alignment between two sequences labeled Qy (query) and Db (database), shown in two columns with paired rows and position numbers.

Left column (positions 961–2040):

Qy 961 CTGCCAGCGTAATCAGTGTATCTATTGTTCCTTCCACTTGCCTACACGAGCCTGGTTG 1020
Db 961 CTGCCAGCGTAATCAGTGTATCTATTGTTCCTTCCACTTGCCTACACGAGCCTGGTTG 1020

Qy 1021 TGTGATCGTCGGGACGAGTGCTGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA 1080
Db 1021 TGTGATCGTCGGCGGGACGAGTGCTGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA 1080

Qy 1081 TTGGACTGGCACGGACGGACTCCTTCCTTGGCTGACCACATTGATTTGTTATGGCGCTCTTGT 1140
Db 1081 TTGGACTGGCACGGACGGACTCCTTCCTTGGCTGACCACATTGATTTGTTATGGCGCTCTTGT 1140

Qy 1141 GACCTGACGCCCTTGACATTGGTGAGTTGTGTGGTGTGTATTAGTTCGGTGACTG 1200
Db 1141 GACCTGACGCCCTTGACATTGGTGAGTTGTGTGGTGTGTATTAGTTCGGTGACTG 1200

Qy 1201 GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGA 1260
Db 1201 GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGA 1260

Qy 1261 AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT 1320
Db 1261 AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT 1320

Qy 1321 CGAGGCTGTGTCATCTTCTTGACCAAACTGGCTTCACAGTACCTATTGCGACTAT 1380
Db 1321 CGAGGCTGTGTCATCTTCTTGACCAAACTGGCTTCACAGTACCATACGCTATTGCGACTAT 1380

Qy 1381 GTTTAGCAGTGTACACTACCTGGCGGTTGGGCGTCTGATCTACTATGCCTCTCGGGCAA 1440
Db 1381 GTTTAGCAGTGTACACTACCTGGCGGTTGGGCGTCTGATCTACTATGCCTCTCGGGCAA 1440

Qy 1441 GTGGTATCAGTTGCTTCCTAGCGCTTATGCTTTACATAGAAGCGACCCTCTGAAACCCAT 1500
Db 1441 GTGGTATCAGTTGCTTCCTAGCGCTTATGCTTTACATAGAAGCGACCCTCTGAAACCTAT 1500

Qy 1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGTGAGTTTGCTCGCCTTTGAGATGATACCATGTCC 1560
Db 1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTCGCCTTTGAGATGATACCATGTCC 1560

Qy 1561 TTGCCACTCTTATTTGTGAGTGAGAATGTGTCAGAAGTCATTTGTTACAGTCCAAAGTGGAC 1620
Db 1561 TTGCCACTCTTATTTGTGAGTGAGAATGTGTCAGAAGTCATTTGTTACAGTCCAAAGTGGAC 1620

Qy 1621 CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGCTACCCCTATACAAATCCCTGG 1680
Db 1621 CAGGCCCTGTCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG 1680

Qy 1681 TGCGAGGGGATGTATGCTACTGACTATGGGCACTGACTGTAAATTCAAAAATAACACATGGG 1740
Db 1681 TGCGAGGGGATGTATGTTGCTACTGACTATGGGCACTGACTGTAAATTCAAAAATAACACATGGG 1740

Qy 1741 TGTGCCATCGACTACTGGAACGACACTCGCAACACTTA 1800
Db 1741 TGTGCCATCGACTATCGTGGAACGACACTCGCAACACTTA 1800

Qy 1801 CGAAGCATGCGGTGTAACACCATGGCTAACAACCGGCTCAGCCTGAA 1860
Db 1801 CGAAGCATGCGGTGTAACACCATGGCTAACAACCGGCTCAGCCTGAA 1860

Qy 1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC 1920
Db 1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC 1920

Qy 1921 AGGCCATTTGTATTTTGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAA 1980
Db 1921 AGGCCATTTGTATTTTGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAA 1980

Qy 1981 TTCCACTCTCCTCCTACCACCCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCTGTGTACG 2040
Db 1981 TTCCACTCTCCTCCTACCACCCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCTGTGTACG 2040

Right column (positions 2041–3121):

Qy 2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100
Db 2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100

Qy 2101 ATTGATACCACCAAAGACAAAGCCTGGAAAAATTATCAGTCTTATATTCGTGGGGTTGTGTGG 2160
Db 2101 ATTGATACCACCAAAGACAAAGCCTGGAAAAATTATCAGTCTTATATTCCGCCACGGGTGC 2160

Qy 2161 TTTGTCTCTTACGGGAGTACCACCAAGGCCGGTGCTAATTCGTGTGGTGTGTGG 2220
Db 2161 TTTGTCTCTTACGGGAGTACCACCAAGGCCGGTGCTAATTCGTGTGGGTGTGTGG 2220

Qy 2221 CAGCAAGTATCTTATTTTAGCCTACCTCGTTACTTGTCCCCTTGTTTTGGGCGCGCTTC 2280
Db 2221 CAGCAAGTATCTTATTTTAGCCTACCTCGTTACTTGTCCCCTTGTTTTGGGGCGCCTTC 2280

Qy 2281 TGGTTACCCCTTGCGTCCGTGCTCCATCCCAAGTCGTATCTCCAAGCTGGCTGGGATGT 2340
Db 2281 TGGTTACCCCTTGCGGTCCGTGCTCCATCCCAAGTCGTATCTCCAAGCTGGCTGGGATGT 2340

Qy 2341 TTTGTCTAAAGCTCAAGTAGCTCTTTGATTTTCTCATCTGTTGCTATCTCCG 2400
Db 2341 TTTGTCTAAAGCTCAAGTAGCTCTTTGATTTTCTCATCTGTTGCTATCTCCG 2400

Qy 2401 CTGCAGGCTACGTTATGCTGCCCCTTTAGGGTTGTGCCCATGGCTGCTGCGGGGCTTGCCCCT 2460
Db 2401 CTGCAGGCTACGTTATGCTGCCCCTTTAGGGTTGTGTGCCCATGGCTGCGGGGCTTGCCCCT 2460

Qy 2461 AACTTCTTGTTGTTGCAGCAGTCTGTGGGCCGGCCGTAACCGTGGAGGCTGGTGGACTGCT 2520
Db 2461 AACTTCTTTGTTGTTGCAGCAGTCTGTGGGCCGGCCGTAACCGTGGAGGCTGGTGGACTGCT 2520

Qy 2521 AGTGGCAGGGTTAGTTTGTGTGGGCCGGTCACCGTAACCGTGTCACCGCTCTGCTTGT 2580
Db 2521 AGTGGCAGGGTTAGTTTGTGTGGGCCGGTCACCGTAACCGTGTCACCGTAGCTCGTTGT 2580

Qy 2581 AGGTCCTTGGCCTCTGTGTAGCGGCTTTAACCCTCCTTGGTTACGCCTGCTTCAGC 2640
Db 2581 AGGTCCTTGGCCTCTGTGTAGCGGCTTTAACCCTCCTTGGTTACGCCTGCTTCAGC 2640

Qy 2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGAGTAGCATTAGTGTCAT 2700
Db 2641 TTTTGACACCGAGATAATTGGAGGGCTGACAATACCACCTGAGTAGCATTAGTGTCAT 2700

Qy 2701 GTCCGTTTGGCTTCTTGCTGTCTTGTTACCTCGCGTGCTTTAGTTAACTCCTATCT 2760
Db 2701 GTCTCGTTTGGCTTCTTGCTGTCTTGTTACCTCGCGTGCTTTAGTTAACTCCTATCT 2760

Qy 2761 TTGGCAACGTTGGGAGAATTGGTTTTCCCGGTGCAACTTACACTCAAGACCGGTGGTACTTTCCT 2820
Db 2761 TTGGCAACGTTGGGAGAATTGGTTTTCCCGGTGCAACTTACACTCAAGACCGGTGGGTTTCTCCT 2820

Qy 2821 TGTGCTGTTGTTTCCCCGGTGGCACATATGACGCGGTGGACTTTCTGTGTGTGTCA 2880
Db 2821 TGTGCTGTTGTTTCCCCGGTGGCACATATGACGCGGTGGACTTTCTGTGTGTGTCA 2880

Qy 2881 CGTAGCTCTTCTATGTTTAACATCCAGTGCAGCAACGGCTCAGCCCTGA 2940
Db 2881 CGTAGCTCTTCTATGTTTAACATCCAGTGCAGCAACGGCTCAGCCCTGA 2940

Qy 2941 TAGGGCCCATAGAATGTTGGTGCGTTCCGGAAAGTGTCATGCTTGGTATTCTCATTATGT 3000
Db 2941 TAGGGCCCATAGAATGTTGGTGCGTTCCGGAAAGTGTCATGCTTGGTATTCTCATTATGT 3000

Qy 3001 TCTTAAGTTTTCCCTCTAGTGTTTGGTGGTGAGAATGTGTTTTTCATAGACCACTTGCA 3060
Db 3001 TCTTAAGTTTTCCCTCTAGTGTTTGGTGGTGAGAATGTGTTTTTCATAGACCACTTGCA 3060

Qy 3061 TGGTGATGTCTTGCCTAATGATTTTGCCTGAAACTACCATTTTCCC 3120
Db 3061 TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTTTTCCC 3120

Qy 3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACGCTTGGCGTGTGGGGACAC 3180

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     TTTTGAAGGCAAGGCAAGGGCTGTATAGAGAATGAAGGAAGACACGCTTGGCGTGTGGGGACAC 3180
Qy GGTTGATGGTTGCCCGTGTTTGCGCGTCTCGGCGACCTTGTTTGCGCGTCTCGGCGACCTTGTTTGCGCAGGGTTGGCTAT 3240
Db GGTTGATGGTTGCCCGTGTTTGCGCGTCTCGGCGACCTTGTTTGCGCAGGGTTAGCTAT 3240
Qy GCCGCCAGATGGGTGGGCCATTACCGCACCTTTTACGCTGTCTTCTGAACGGTGG 3300
Db GCCGCCCAGATGGGTGGGCCATTACCGCACCTTTTACGCTGTCTTCTGAACGGTGG 3300
Qy CACGCTGTCAGCGATGGCAGTGCAGTGTCATGACTGTATAGACCCCGAACTTGGACTGGAAC 3360
Db CACGCTGTCAGCGATGGCAGTGCAGTGTCATGACTGTATAGACCCCGAACTTGGACTGGAAC 3360
Qy TATCTTCAGATTAGGAATCTCTGGCCACTAGCTACATGGGATTTGTTGTTGTGACAACGTGTT 3420
Db TATCTTCAGATTAGGAATCTCTGGCCACTAGCTACATGGGATTTGTTGTTGTGACAACGTGTT 3420
Qy GTATACTGCTCACCATGCAGCAAGGGCGCCGGTTGCTCATCCCACAGGCTCTATACA 3480
Db GTATACTGCTCACCATGCAGCAAGGGCGCCGGTTGCTCATCCCACAGGCTCCATACA 3480
Qy CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCACCATGTGGAGCGTGG 3540
Db CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCACCATGTGGAGCGTGG 3540
Qy GTCCCTTACTCGGTGCTCTTGCGCGGGAGGACCAAGGGTATCGGTAACACGGACTGGGGTC 3600
Db GTCCCTTACTCGGTGCTCTTGCGCGGGAGGACCAAGGGTATCGGTAACACGGACTGGGGTC 3600
Qy ATTGGGTTGAGGTCAACAAATCCGATGACCCCTTATTGGTGTGTGTGCGGGGCCCTTCCCAT 3660
Db ATTGGGTTGAGGTCAACAAATCCGATGACCCCTTATTGGTGTGTGTGCGGGGCCCTTCCCAT 3660
Qy GGCTGTTGCCAAGGGTTCTTCAGGTGCCCGATTCTGTGCTCCCGGGCATGTTATTGG 3720
Db GGCTGTTGCCAAGGGTTCTTCAGGTGCCCGATTCTGTGCTCCCGGGCATGTTATTGG 3720
Qy GATGTTCACCGTGCTAGAAATTCTGGCGGTTCAGTCAGTAGGGTTAGGCCGTT 3780
Db GATGTTCACCGTGCTAGAAATTCTGGCGGTTCAGTCAGTCAGCCAGATTAGGGTTAGGCCGTT 3780
Qy GGTGTGTGCTGGATACCATCCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC 3840
Db GGTGTGTGCTGGATACCATCCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC 3840
Qy TGTGCCTAACGAGTATTCGTTAAATTGCCCCCACTGGCCAGCGGGCAAGTCAAC 3900
Db TGTGCCTAACGAGTATTCGTTAAATTGCCCCCACTGGCAGCGGGCAAGTCAAC 3900
Qy CAAAATTACCACTTTCTTACATGCAGGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT 3960
Db CAAAATTACCACTTTCTTACATGCAGGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT 3960
Qy GGCTACAACAGCATCAATGCCAAAGTACGCACGACGTACGGCCTGCCACTTTGTGACGAATTG 4020
Db GGCTACAACAGCATCAATGCCAAAGTACGCACGACGTACGGCCTGCCACTTTGTGACGAAATTG 4020
Qy CTATTTTAATGCAAATGTGTTGGCATTGGCAAAGTCAACGTGGCATTGGGCATCCAAATGGCAT 4080
Db CTATTTTAATGCAAATGTGTTGGCATTGGCAAAGTCAACGTGGCATTGGGCATCCAAATGGCAT 4080
Qy GTACCTGACCGGAGCCATGTTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGC 4140
Db GTACCTGACCGGAGCCATGTTCCCGGAACTATGATGTCATCATTTGTGACGAATGCCATGC 4140
Qy TACCGATGCAACCACCGTGTTGCCGCACCGTGTTCTTGGAAAGGTTCCATCCAAAA 4200
Db TACCGATGCAACCACCGTGTTGCCGCACCGTGTTCTTGGAAAGGTTCCATCCAAAAA 4200
Qy TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCTGGAGTAATCCTACACCACATGC 4260
Db TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCTGGAGTAATCCTACACCACATGC 4260

Db TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCTGGAGTAATCCTACACCACATGC 4260
Qy CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTCATGGAAAAAGAT 4320
Db CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTCATGGAAAAAAGAT 4320
Qy TAAGGAGGAAAATCTGAAGAAAGGGAGACACCTTATCTTGAGGCTACCAAAAAACACTG 4380
Db TAAGGAGGAAAATCTGAAGAAAGGGAGACACCTTATCTTGAGGCTACCAAAAAACACTG 4380
Qy TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440
Db TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440
Qy ATGTGACATCTCAAAAAATCCTGAGGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500
Db ATGTGACATCTCAAAAAATCCTGAGGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500
Qy TACAGGGTACACTCGTGACTTTGATTCCGTGATGACTGGTGTATGACTGTAGAAGGCAC 4560
Db TACAGGGTACACTCGTGACTTTGATTCCGTGATGACTGGTGTATGACTGTAGAAGGCAC 4560
Qy ATGCCATGTGACCTTGACCCTACTTTGCGTGTGTGTGTGTGTCGTTCAGC 4620
Db ATGCCATGTGACCTTGACCCTACTTTGCGTGTGTGTGTGTGTCGTTCGTTCTCAGC 4620
Qy AATAGTTAAAGGCCAGCGTGAGGGGCCGCACAGGCCGTGGGAGAGCTGGCGCATATACTACTA 4680
Db AATAGTTAAAGGCCAGCGTGAGGGGCCGCACAGGCCGTGGGAGAGCTGGCGCATATACTACTA 4680
Qy TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740
Db TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740
Qy CGACGCAGCCAAGGCATGTATGGTATGTCATCAACAGAAGCTCAAACTATTCTGGACAC 4800
Db CGACGCAGCCAAGGCATGTATGGTATGTCATCAACAACAAGCTCAAACTATTCTGGACAC 4800
Qy CTATCGCACCACCAACCTGGCAGGAGCAAATTTGACGAGTGGGGCTGATCT 4860
Db CTATCGCACCACCAACCTGGCAGGAGCAAATTTGACGAGTGGGGCTGATCT 4860
Qy CTTTTCTATGGTCAACCCGAACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA 4920
Db CTTTTCTATGGTCAACCCGAACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA 4920
Qy TTATGTTTGTTGACTGAGCAGCCCAACTACAACTGTGTCATCAGTATGGTCATGCTGCTCC 4980
Db TTATGTTTGTTGACTGAGCAGCCCAACTACAACTGTGTCATCAGTATGGTCATGCTGCTCC 4980
Qy CAATGACGCACCACGGCTGCAGGAGCGCCGCCCGGCTTGTGGGAAAAACCTTGTGGGGTTCTGTG 5040
Db CAATGACGCACCACGGCTGCAGGAGCGCCGCCCGGCTTGTGGGAAAAAACCTTGTGGGGTTCTGTG 5040
Qy GCGCTTGGACGGCGGCTGACGCCTTCTTGCCCAGAGCCCAGGAGGTGACCAGGATACCA 5100
Db GCGCTTGGACGGCGGCTGACGCCTTCTTGCCCAGAGCCCAGGAGGTGACCAGATACCA 5100
Qy AATTGCTTCACTGAAGTCAATACTTCTGGACAGCCCGGCACTCCGCTCGTGTGTGGTGGGAGT 5160
Db AATGTGCTTCACTGAAGTCAATACTTCTGGACAGCCCGGCACTCCGCTCGTGTGTGGTGGGAGT 5160
Qy GGCTATGGCTTATCTAGCCATTGGGAACTATCAGTATCATTTGGCGGTATGACGAATGCCATGC 5220
Db GGCTATGGCTTATCTAGCCATTGGGAACTATCAGTATCATTTGGCGGTATGACGAATGCCATGC 5220
Qy TATTACAATCAGTCCCTACCGGTGGTTCACTTGGCGGTTGTGGCGGTTGCGGTTGCGCAGAAATCGT 5280
Db TATTACAATCAGTCCCTACCGGTGGTTCACTTGGCGGTTGTGGCGGTTGCGGTTGCGCAGAAATCGT 5280
Qy GGAGGAGTGTGCAATCATTCATTCCCTTGGAGGCCATGGGTGCAATCGATAAGCTGAA 5340
Db GGAGGAGTGTGCAATCATTCATTCCCTTGGAGGCCATGGGTGCAATCGATAAGCTGAA 5340

The body of this page is a dense two-column DNA sequence alignment consisting of paired "Qy" (query) and "Db" (database) rows with match bars, numbered by base position. Left column spans positions 5341–6420; right column spans positions 6421–7560.

```
Db  7501 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 7560
Qy  7561 CCACATCACTGGCCTTCGGGCGGCCACTGATGTCGTTCGTTCTCGGAGCAGCCCGCAAGGCTGTTCT 7620
Db  7561 CCACATCACTGGCCTTCGGGCGGCCACTGATGTCGTTCGTTCTCGGAGCAGCCCGCAAGGCTGTTCT 7620
Qy  7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680
Db  7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680
Qy  7681 AGTTCCAAAGGAGGAGGTCTTCGTGAGAACCAACAACAAGAAACCCCCAAG 7740
Db  7681 AGTTCCAAAGGAGGAGGTCTTCGTGAGAACCAACAACAAGAAACCCCCAAG 7740
Qy  7741 GCTTATCTCGTACCCCACCTTGAAATCAGATGTGTTGAGAAGATGTACTACGGTCAGGT 7800
Db  7741 GCTTATCTCGTACCCCACCTTGAAATCAGATGTGTTGAGAAGATGTACTACGGTCAGGT 7800
Qy  7801 TGCTCCTGACGTAGTTAAAGCTGTCTGTGGAGATGCGTACGGGTTTGTAGATCCACGTAC 7860
Db  7801 TGCTCCTGACGTAGTTAAAGCTGTCTGTGGAGATGCGTACGGGTTTGTAGATCCACGTAC 7860
Qy  7861 CCGTGTCAAGGCTGTCTGTGTTGTCGATGTGGTCACCCGAGCAGTCGGAGGCCACATGCCACGATAC 7920
Db  7861 CCGTGTCAAGGCTGTCTGTGTTGTCGATGTGGTCACCCGAGCAGTCGGAGGCCACATGCCACGATAC 7920
Qy  7921 AGTGTGTTTGACAGTACCATCACCAACACCCGAGGATATCATGTGGTGGACACAGACAGCATCTACTC 7980
Db  7921 AGTGTGTTTGACAGTACCATCACCAACACCCGAGGATATCATGTGGTGGACACAGACAGCATCTACTC 7980
Qy  7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTAGGTC 8040
Db  7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTAGGTC 8040
Qy  8041 CGCTGGAGGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTAGGTC 8100
Db  8041 CGCTGGAGGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTAGGTC 8100
Qy  8101 TTCCGGGGTCTATACTCCAAGTTCCAACAGTTTGACCTGCTGCTGGCTGAAGGTAAATGC 8160
Db  8101 TTCCGGGGTCTATACTCCAAGTTCCAACAGTTTGACCTGCTGCTGGCTGAAGGTAAATGC 8160
Qy  8161 TGCAGCCGAACAGGCTGCATGAGAACCCCTCGCTTCCTTATTTGCGGCGATGATTGCAC 8220
Db  8161 TGCAGCCGAACAGGCTGCATGAGAACCCCTCGCTTCCTTATTTGCGGCGATGATTGCAC 8220
Qy  8221 CGTAATTTGGAAGAGCGCCGGAGCGCGGAGCAGAACAAGCAATGCGTGTCTTGCTAG 8280
Db  8221 CGTAATTTGGAAGAGCGCCGGAGCGCGGAGCAGAACAAGCAATGCGTGTCTTGCTAG 8280
Qy  8281 CTGGATGAAGGTGATGGGTGCACCACCAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340
Db  8281 CTGGATGAAGGTGATGGGTGCACCACCAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340
Qy  8341 AGAATTAACATCATGCTGCCTGGAATTACCTCTGGAAGTGGCAAGCCTTA 8400
Db  8341 AGAATTAACATCATGCTGCCTGGAATTACCTCTGGAAGTGGCAAGCCTTA 8400
Qy  8401 CTACTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTGCCGAGGGTCTGGG 8460
Db  8401 CTACTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTGCCGAGGGTCTGGG 8460
Qy  8461 ATACAACCCCAGTGCTGCGTGTCCATTTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8520
Db  8461 ATACAACCCCAGTGCTGCGTGTCCATTTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8520
Qy  8521 TAGCCGTGTTGGCTGTGTCCATTTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580
Db  8521 TAGCCGTGTTGGCTGTGTCCATTTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580
Qy  8581 GACTGTGACCTTTGACTGTGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640
```

```
Db  8581 GACTGTGACCTTTGACTGTGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640
Qy  8641 CATCATTGCTGGTGTGCACGGTATTGAGGCTTTCTCGGTGTCGGCGCTACCACCAACGCTGA 8700
Db  8641 CATCATTGCTGGTGTGCACGGTATTGAGGCTTTCTCGGTGTCGGCGCTACCACCAACGCTGA 8700
Qy  8701 GATCCTCAGAGTTCCAATCACTCAACAGACATGACCATGACCATGCCCCCTGCGAGCCTGGCG 8760
Db  8701 GATCCTCAGAGTTCCAATCACTCAACAGACATGACCATGACCATGCCCCCTGCGAGCCTGGCG 8760
Qy  8761 AAAGAAAGCCAGGGCGGTCCTGCCAGCGCCAGCGCCAAGAGGCGTGGGCGGAGCACACGCAAAATT 8820
Db  8761 AAAGAAAGCCAGGGCGGTCCTGCCAGCGCCAGCGCCAAGAGGCGTGGGCGGAGCACACGCAAAATT 8820
Qy  8821 GGCTGCGCTTCCTTCCTCTCGGCATGCTACATCTAGACCTCTACCAGATTTGGATAAGACGAG 8880
Db  8821 GGCTGCGCTTCCTTCCTCTCGGCATGCTACATCTAGACCTCTACCAGATTTGGATAAGACGAG 8880
Qy  8881 CGTGGCTCGGTACACCACTTCAATTATTGTGATGTTTACTCCCCGGAGGGGGATGTGTT 8940
Db  8881 CGTGGCTCGGTACACCACTTCAATTATTGTGATGTTTACTCCCCGGAGGGGGATGTGTT 8940
Qy  8941 TATTACACCACAGAGAAGATTGCAGAAGTTCCTTGTGAAGTATTTGGGCTGTCATTGTTTT 9000
Db  8941 TGTTACACCACAGAGAAGATTGCAGAAGTTCGTGTGAAGTATTTGGCTGTCATTGTTTT 9000
Qy  9001 TGCCCTAGGGCTCATTGCTGTGTTGGACTAGCCATCAGTGACTGAACCCCAAATTCAAAATTAA 9060
Db  9001 TGCCCTAGGGCTCATTGCTGTGTTGGACTAGCCATCAGTGACTGAACCCCAAATTCAAAATTAA 9060
Qy  9061 CTAACAG---TTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCAGCAACAGGGGAGACCCC 9116
Db  9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCAGCAACAGGGGAGACCCC 9120
Qy  9117 GGGCTTAACGACCCCGC 9133
Db  9121 GGGCTTAACGACCCCGC 9137
```

RESULT 3
US-08-469-260A-393
; Sequence 393, Application US/08469260A
; Patent No. 6451578
GENERAL INFORMATION:
APPLICANT: JOHN N. SIMONS
APPLICANT: TAMI J. PILOT-MATIAS
APPLICANT: GEORGE J. DAWSON
APPLICANT: GEORGE G. SCHLAUDER
APPLICANT: SURESH M. DESAI
APPLICANT: THOMAS P. LEARY
APPLICANT: ANTHONY SCOTT MUERHOFF
APPLICANT: JAMES C. ERKER
APPLICANT: SHERI L. BUIJK
APPLICANT: ISA K. MUSHAHWAR
TITLE OF INVENTION: NON-A, NON-B. NON-C, NON-D, NON-E HEPATITIS
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
NUMBER OF SEQUENCES: 716
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,260A
FILING DATE:
CLASSIFICATION:

```
;  PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US/08/424,550
;      FILING DATE:
;  ATTORNEY/AGENT INFORMATION:
;      NAME: POREMBSKI, PRISCILLA E.
;      REGISTRATION NUMBER: 33,207
;      REFERENCE/DOCKET NUMBER: 5527.PC.01
;  TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 708-937-6365
;      TELEFAX: 708-938-2623
;  INFORMATION FOR SEQ ID NO: 393:
;      SEQUENCE CHARACTERISTICS:
;          LENGTH: 9143 base pairs
;          TYPE: nucleic acid
;          STRANDEDNESS: double
;          TOPOLOGY: linear
;      MOLECULE TYPE: DNA (genomic)
;      FEATURE:
;          NAME/KEY: 5'UTR
;          LOCATION: 1..445
;      FEATURE:
;          NAME/KEY: CDS
;          LOCATION: 446..9037
;      FEATURE:
;          NAME/KEY: 3'UTR
;          LOCATION: 9038..9143
US-08-469-260A-393

Query Match            96.4%;   Score 9059.8;  DB 3;  Length 9143;
Best Local Similarity  99.6%;   Pred. No. 0;
Matches 9096;  Conservative  0;  Mismatches 37;  Indels  4;  Gaps  1;
```

```
Qy    1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGAGCACACCCCCCCTAG   60
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGAGCACACCCCCCCTAG   60

Qy   61 CAGGGCGTGGGGATTCCCCGCCTCTGCAGAAGGGTGGAGCCAACCACCTTAGTAT      120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CAGGGCGTGGGGATTCCCCGCCTCTGCAGAAGGGTGGAGCCAACCACCTTAGTAT      120

Qy  121 GTAGGGCGCGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTAGGGCGCGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180

Qy  181 CCTGATGGGCGGTTCATGGGTTCGGTGTGGTGGTCGGTGGTGGTGGGCGCTTTAGGCAGC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CCTGATGGGCGGTTCATGGGTTCGGTGTGGTGGTCGGTGGTGGTGGGCGCTTTAGGCAGC  240

Qy  241 CCTCCCAGATAGAGCGGCGGCGGCCACTGTAGGGGAAGAACCGGGGACCCGTCACTACCAAGGACG  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 CCTCCCAGATAGAGCGGCGGCGGCCACTGTAGGGGAAGAACCGGGGACCCGTCACTACCAAGGACG  300

Qy  301 CAGACCTCTTTTGAGTATCACGCCCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTGT  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGACCTCTTTTGAGTATCACGCCCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTGT  360

Qy  361 TGGGATGGTTGGGGTTAGCCATCCATCCATCCATCCGTACTACGTCCTTGCGAGGGGAT  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 TGGGATGGTTGGGGTTAGCCATCCATCCATCCATCCGTACTACGTCCTTGCGAGGGGAT  420

Qy  421 CTGGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACC  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CTGGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACC  480

Qy  481 TGCGCCCAGAAGACGCGCAAGAACAAGCAGAGACACGCGCAAGCGCAAAGTCTGTCATTAAAAC  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TGCGCCCAGAAGACGCGCAAGAACAAGCAGAGACACGCGCAAGCGCAAAGTCTGTCATTAAAAC  540

Qy  541 ATCTGTTGAAAGGGGACAACGACAAAGCGCAAAGTCCAGCCGACGCTCCGCCCTCGTAA  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 ATCTGTTGAAAGGGGACAACGACAAAGCGCAAAGTCCAGCCGACGCTCCGCCCTCGTAA  600

Qy  601 TTACAAAATTGCTCGGTATCCATGATGCCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 TTACAAAATTGCTCGGTATCCATGATGCCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC  660

Qy  661 TCATGGTTGGGGACGGCCAAGACCCTGCGCAATCTTGGAATCCTTCTGA            720
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TCATGGTTGGGGACGGCCAAGACCCTGCGCAATCTTGGAATCCTTCTGA            720

Qy  721 TTACCCTTTGGGGTGGATTGGTGATGTTACAACTCACACACCTTAGTAGGCCCGCTCGT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 TTACCCTTTGGGGTGGATTGGTGATGTTACAACTCACACACCTTAGTAGGCCCGCTCGT  780

Qy  781 GGCAGGAGCGGTCGTTGACCAGTCTCGCAGATAGTACGCTTGCTGGGAGGATGGAGTCAA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GGCAGGAGCGGTCGTTGACCAGTCTCGCAGATAGTACGCTTGCTGGGAGGATGGAGTCAA  840

Qy  841 CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTGTGGTATGTCTGCTATCTTGGCCTG    900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTGTGGTATGTCTGCTATCTTGGCCTG    900

Qy  901 TCCCTGTAGTGGGGCGCGGGGTCACTGACCCAGACACAAATACCACACAATCCTGACCAATTG  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TCCCTGTAGTGGGGCGCGGGGTCACTGACCCAGACACAAATACCACACAATCCTGACCAATTG  960

Qy  961 CTGCCCAGCTAATCAGGTTATCTATTGTCTCCTTCCACTTGCTACACGAGCCTGGTTG     1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 CTGCCCAGCTAATCAGGTTATCTATTGTCTCCTTCCACTTGCTACACGAGCCTGGTTG     1020

Qy 1021 TGTGATCTGTCCGGACGAGTGCTGGGGTTCCCGCCAATCCGTACATCTCACACCCTTCAA  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 TGTGATCTGTCCGGACGAGTGCTGGGGTTCCCGCCAATCCGTACATCTCACACCCTTCAA  1080

Qy 1081 TTGGACTGGCACGGACTCCTTCTTGGCTGACCACATTGATTTTGTTATGGGCGCTCTTGT  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 TTGGACTGGCACGGACTCCTTCTTGGCTGACCACATTGATTTTGTTATGGGCGCTCTTGT  1140

Qy 1141 GACCTGTGAGCGCCCTTGACATTGTGAGTTGTCGTGTGTATTAGTCGGTGACTG          1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 GACCTGTGAGCGCCCTTGACATTGTGAGTTGTCGTGTGTATTAGTCGGTGACTG          1200

Qy 1201 GCTTGTCAGGCACTGGCTTAATCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 GCTTGTCAGGCACTGGCTTAATCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGA  1260

Qy 1261 AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGCCGTGGCCGCAAGGT  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGCCGTGGCCGCAAGGT  1320

Qy 1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT  1380

Qy 1381 GTTTAGCAGTGTACACTACCACCTGGCGGTTGGCGCTGCGGTGGGCGGTCTCGTGGGGCAA  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 GTTTAGCAGTGTACACTACCACCTGGCGGTTGGCGCTGCGGTGGGCGGTCTCGTGGGGCAA  1440

Qy 1441 GTGGTATCAGTTGCTCCTCCTAGCCGCTTACATAGAAGGGACCTCTGGAAACCCCAT      1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 GTGGTATCAGTTGCTCCTCCTAGCCGCTTACATAGAAGGGACCTCTGGAAACCCCAT      1500

Qy 1501 CAGGGTGCCCACTGGGATGCTCAATAGCTGAGTTTTCGCCCTTTGATGATACCATGTCC  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 CAGGGTGCCCACTGGGATGCTCAATAGCTGAGTTTTCGCCCTTTGATGATACCATGTCC  1560

Qy 1561 TTGCCACTCTTATTTGAGTGAAATGTGTCGAAGTCATTGTTACAGTCCAAAGTGGAC  1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 TTGCCACTCTTATTTGAGTGAAATGTGTCGAAGTCATTGTTACAGTCCAAAGTGGAC  1620

Qy 1621 CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCTGG   1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG   1680

Qy 1681 TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA  1740
```

Db   1681 TGCGAGGGGATGGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA 1740

Qy   1741 TGTGCCATGCGTACTGCACTGACTATGGGCACTGGGAACGACACTCGCAACACTTA   1800
            ||||||||||||| ||||||||||||||||||||| ||||||||||||||||||||
Db   1741 TGTGCCATGCGTACTGCACTGACTATGGGCACTGTGGAACGACACTCGCAACACTTA   1800

Qy   1801 CGAAGCATGCGGTGTAACACCATGGCTAACACGGCACATGGCACAACGGCTCAGCCCTGAA 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CGAAGCATGCGGTGTAACACCATGGCTAACACGGCACATGGCACAACGGCTCAGCCCTGAA 1860

Qy   1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC   1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC   1920

Qy   1921 AGGCCATTTGTATTTTGAGGGATCAGATACCCTATAGTTTACTTTTTATGACCCTGTGAA 1980
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AGGCCATTTGTATTTTGAGGGATCAGATACCCTATAGTTTACTTTTTATGACCCTGTGAA 1980

Qy   1981 TTCCACTCTCCTACCACCGGAGAGGTGGGCTAGTTGCCCGGTACCCGGTACGGTACG   2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TTCCACTCTCCTACCACCGGAGAGGTGGGCTAGTTGCCCGGTACCCGGTACGGTACG   2040

Qy   2041 TGGTTCTTGGTTTACAGGTTCCGCAAGGTTTTACAGTGATGTGAAAGACCTACCACAGG   2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TGGTTCTTGGTTTACAGGTTCCGCAAGGTTTTACAGTGATGTGAAAGACCTACCACAGG   2100

Qy   2101 ATTGATCACCAAAGACAAAGCCTGGAAAAAATTATCAGTTCTTATATTCCGCCACGGGTGC 2160
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ATTGATCACCAAAGACAAAGCCTGGAAAAAATTATCAGTTCTTATATTCCGCCACGGGTGC 2160

Qy   2161 TTTGTCTTCCTTACGGGAGTTACCACCAAGGCCGTGTGTCTAATTCGTGTGTGTTTCGTGTGG 2220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TTTGTCTTCCTTACGGGAGTTACCACCAAGGCCGTGTGTCTAATTCGTGTGTGTTTCGTGTGG 2220

Qy   2221 CAGCAAGTATCTTATTTTAGCTACCTCTGTTACTTGTCCCTTGTTTGTTTGGGCGCGCTTC 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CAGCAAGTATCTTATTTTAGCTACCTCTGTTACTTGTCCCTTGTTTGTTTGGGCGCGCTTC 2280

Qy   2281 TGGTTACCCTTTGCGTCCTGTGCTCCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGT 2340
            ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 TGGTTACCCTTTGCGTCCTGTGCTCCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGT 2340

Qy   2341 TTTGTCTAAAGCTCAAGTAGCTTCCTTTGATTTTCTTCATCGTGTCTATCTCCG   2400
            |||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   2341 TTTGTCTAAAGCTCAAGTAGCTTCCTTTGCTTTGATTTTCTTCATCGTGTCTATCTCCG   2400

Qy   2401 CTGCAGGCTACGTTATGCTGCCCTTTAGGGTTTGTGCCCATGGCTTGCGGGCTTGCCCCT 2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 CTGCAGGCTACGTTATGCTGCCCTTTAGGGTTTGTGCCCATGGCTTGCGGGCTTGCCCCT 2460

Qy   2461 AACTTCTTGTTGCAGCAGCGTCTGCGCCAACCAGAATATGACTATGCTAGTTAGTTGTCAT 2520
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 AACTTCTTGTTGCAGCAGCGTCTGCGCCAACCAGAATATGACTATGCTAGTTAGTTGTCAT 2520

Qy   2521 AGTGGCAGGGTAGTTTGTTTGTGGGCGGCCGCCGTAACCGTACGTGTCACGTAGCTCTGTT 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 AGTGGCAGGGTAGTTTGTTTGTGTGGGCGGCCGCCGTGACCGTACGTGTCACGTAGCTCTGTT 2580

Qy   2581 AGTCCTTGGCCTCTCGGTAGCGTCTTTAACCCTCTTGGTTAACGCCTTGCTTCAGC   2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 AGTCCTTGGCCTCTCGGTAGCGTCTTTAACCCTCTTGGTTAACGCCTTGCTTCAGC   2640

Qy   2641 TTTTGATACCGAGATAATTGGAGGGGCTAGTAGCACCACCTGTAGTAGCATTAGTTGTCAT 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 TTTTGACACCGAGATAATTGGAGGGGCTAGTAGCACCACCTGTAGTAGCATTAGTTGTCAT 2700

Qy   2701 GTCTCGTTTGGCCTCTCGGTAGCGTCTCTGTCTGTGCTCTGTGCTTAGTTAACTCCTATCT 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GTCTCGTTTGGCCTCTCGGTAGCGTCTCTGTCTGTGCTCTGTGCTTTAGTTAACTCCTATCT 2760

Qy   2761 TTTGGCAACGTTGGGAGAATTGGTTTTTGGAACGTTACACTAAGACCGGAGAGGTTTCTCCT 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 TTTGGCAACGTTGGGAGAATTGGTTTTTGGAACGTTACACTAAGACCGGAGAGGTTTCTCCT 2820

Qy   2821 TGTGCTGGTTGTTGTTCCCGGTGCGACATATGACGCGCTGGTCGACTTTCTGTGTGTGTCA 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 TGTGCTGGTTTGTTGTTCCCGGTGCGACACATGACGCGCTGGTCGACTTTCTGTGTGTGTCA 2880

Qy   2881 CGTAGCTCTCATGTTTAACATCCAGTGCAGCATCGTTCTTTGGGACTGACTTCTAGGGT   2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 CGTAGCTCTTCATGTTTAACATCCAGTGCAGCATCGTTCTTTGGGACTGACTTCTAGGGT   2940

Qy   2941 TAGGGCCCATAGAAGTGTTGGTGCGTCTCCGAAAGTGTCATGCTTGGTATTCTCATTATGT 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 TAGGGCCCATAGAAGTGTTGGTGCGTCTCCGGAAAGTGTCATGCTTGGTATTCTCATTATGT 3000

Qy   3001 TCTTAAGTTTTTCCTCTTAGTGTTTGGTGAGAATGGTGTTTTCTATAAGCACTTGCA   3060
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 TCTTAAGTTTTTCCTCTTAGTGTTTGGTGAGAATGGTGTTTTCTATAAGCACTTGCA   3060

Qy   3061 TGGTGATGCTCTTGCCTAATGATTTTGCCTGGAAACTACCATTGCAAGAGCCATTTTTCCC 3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 TGGTGATGCTCTTGCCTAATGATTTTGCCTGGAAACTACCATTGCAAGAGCCATTTTTCCC 3120

Qy   3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACGCTTGGCGTGTGGGGACAC 3180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACGCTTGGCGTGTGGGGACAC 3180

Qy   3181 GGTTGATGGTTTGCCCGTTGTTGCCCGTCTCGGCGACCTTGTTTTCGCAGGGTTGGCTAT 3240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 GGTTGATGGTTTGCCCGTTGTTGCCCGTCTCGGCGACCTTGTTTTCGCAGGGTTAGCTAT 3240

Qy   3241 GCCCCAGATGGTGGGCCATTACCGCCACCTTTTACGCTGCAGTGTCTCTGCAGTGTCGTCGG 3300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 GCCCCAGATGGGTGGGCCATTACCGCCACCTTTTACGCTGCAGTGTCTCTCGAACGTCGG 3300

Qy   3301 CACCCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGAGACCCCGAACTTGGACTGGAAC 3360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGAGACCCCGAACTTGGACTGGAAC 3360

Qy   3361 TATCTTCAGATTAGGATCTCTGCCACTAGCTCAACATGGGATTTGTTTGTGACAACGTGTT 3420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 TATCTTCAGATTAGGATCTCTGCCACTAGCTCAACATGGGATTTGTTGTGTGACAACGTGTT 3420

Qy   3421 GTATACTGCTCACCATGCAGCAAGGGCGCCGGTTGGCTCATCCCACCAGGCTCTCTATACA 3480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 GTATACTGCTCACCATGCAGCAAGGGCGCCGGTTGGCTCATCCCACCAGGCTCTCATACA 3480

Qy   3481 CCCAATAACCGTTGACGCGGCTAATGACCGGAGACATCTATCAACCACCATGTGGAGCTGG 3540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 CCCAATAACCGTTGACGCGGCTAATGACCGGAGACATCTATCAACCACCATGTGGAGCTGG 3540

Qy   3541 GTCCCTTACTCGGTGCTCTTGCGGGGAGACCAAGGGTATCTGGTAACGACTGGGGTC   3600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 GTCCCTTACTCGGTGCTCTTGCGGGGAGACCAAGGGTATCTGGTAACGACTGGGGTC   3600

Qy   3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGCGGGGCCCTTCCCAT   3660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGCGGGGCCCTTCCCAT   3660

Qy   3661 GGCTGTTGCCAAGGGTTCTTCAGGTGCGCCGGGCCCCGATTCGTGTGCTCCTCCCGGGCATGTTATTGG 3720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 GGCTGTTGCCAAGGGTTCTTCAGGTGCGCCGGGCCCCGATTCGTGTGCTCCTCCCGGGCATGTTATTGG 3720

Qy   3721 GATGTTCACCGCTGCTAGAAATTCTGGCGGTTCAGTCAGTTAGGGTTAGGGTTAGGCCGTT 3780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 GATGTTCACCGCTGCTAGAAATTCTGGCGGTTCAGTCAGTTAGGGTTAGGGTTAGGCCGTT 3780

Qy   3781 GGTGTGTGCTGGATACCATCCCCAGTACACAGCGCCACTCTTGATACAAAAACCTAC   3840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 GGTGTGTGCTGGATACCATCCCCAGTACACAGCACACATGCCACTCTTGATACAAAAACCTAC 3840

Qy   3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC   3900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC   3900

Left column:

```
Qy  3901 CAAATTACCACTTCTTACATGCAGGAGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT 3960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 CAAATTACCACTTCTTACATGCAGGAGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT 3960

Qy  3961 GGCTACAACAGCATCAATGCCAAAGTACATGCGACGTACGGCGTACGGCGAATCAAATTG 4020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 GGCTACAACAGCATCAATGCCAAAGTACATGCGACGTACGGCGTACGGCGAATCAAATTG 4020

Qy  4021 CTATTTTAATGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGCAT 4080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 CTATTTTAATGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGCAT 4080

Qy  4081 GTACCTGACCGGAGCATGTTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGC 4140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 GTACCTGACCGGAGCATGTTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGC 4140

Qy  4141 TACCGATCGCAACCACCGTGTTGGGCATTGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA 4200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 TACCGATCGCAACCACCGTGTTGGGCATTGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA 4200

Qy  4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCGGAGTAATCCCTTTCATGGAACCACATGC 4260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCGGAGTAATCCCTTTCATGGAACCACATGC 4260

Qy  4261 CAACATAACTGAGATTCAATTAACCGATGAAGGGCACTCCTATCCCTTTCATGGAACACACTG 4320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261 CAACATAACTGAGATTCAATTAACCGATGAAGGGCACTCCTATCCCTTTCATGGAACACACTG 4320

Qy  4321 TAAGGAGGAAAATCTGAAGAAAGGGAGACACCTTATCTTTGAGGCTACCAAAAAAAACACTG 4380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321 TAAGGAGGAAAATCTGAAGAAAGGGAGACACCTTATCTTTGAGGCTACCAAAAAAAACACTG 4380

Qy  4381 TGATGAGCTTGCTACCAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381 TGATGAGCTTGCTACCAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440

Qy  4441 ATGTGACATCTCAAAAATCCAAAAAATCCCGAGGGCGCACTGGGTGCCACTGATGCCTGTG 4500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441 ATGTGACATCTCAAAAATCCAAAAAATCCCGAGGGCGCACTGGGTGCCACTGATGCCTGTG 4500

Qy  4501 TACAGGGTACACTGGTGACTTGACTTTGATTCCGTGTATGACTGCAGCCTCATGTGATAGAAGGCAC 4560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501 TACAGGGTACACTGGTGACTTGACTTTGATTCCGTGTATGACTGCAGCCTCATGTGATAGAAGGCAC 4560

Qy  4561 ATGCCATGTTGACCTTGACCCTACTTTCACCATGCCATGGTGTCGTGTGCGGGTTTCAGC 4620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561 ATGCCATGTTGACCTTGACCCTACTTTCACCATGCCATGGTGTCGTGTGCGGGTTTCAGC 4620

Qy  4621 AATAGTTAAAGGCCAGCCGTGGGGTTACCTGGGGAGAGCGGTTGGCATATACTACTA 4680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621 AATAGTTAAAGGCCAGCCGTGGGGTTACCTGGGGAGAGCGGTTGGCATATACTACTA 4680

Qy  4681 TGTAGAGACGGGAGTTGTACCCCGTATGGTTCCTCAACATTGTTGAAGCCTT 4740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681 TGTAGAGACGGGAGTTGTACCCCGTATGGTTCCTCAACATTGTTGAAGCCTT 4740

Qy  4741 TTATGTTTTGTTGTTAGTCAGCCGAGCATGGTGTGTCATCATGTCTATGGCTATGCTGCTCC 4800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741 TTATGTTTTGTTGTTAGTCAGCCGAGCATGGTGTGTCATCATGTCTATGGCTATGCTGCTCC 4800

Qy  4801 CGACGCAGCCAAGGCCAAGGCATGGTATGGTTTGTCATCAACAACAGAAGCTCAAACTAATTCTGACACAC 4860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801 CGACGCAGCCAAGGCCAAGGCATGGTATGGTTTGTCATCAACAACAGAAGCTCAAACTAATTCTGACACAC 4860

Qy  4861 CTTTTCTATGGTCAACCCGAACCTTCATTTGTCAATACTGTCAAAAGAACTGCTGACAA 4920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861 CTTTTCTATGGTCAACCCGAACCTTCATTTGTCAATACTGTCAAAAGAACTGCTGACAA 4920

Qy  4921 TTATGTTTGTTGTTACTGACTGCAGCCAACTACAACTGTCATCAGTATGGCTATGCTGCTCC 4980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921 TTATGTTTGTTGTTACTGACTGCAGCCAACTACAACTGTCATCAGTATGGCTATGCTGCTCC 4980

Qy  4981 CAATGACGCACCACCGGTGGGCAGGGAGCCCGGCTTGTGGGAAAAAAACCTTGTGGGGGTTCTGTG 5040
```

Right column:

```
Db  4981 CAATGACGCACCACCGGTGGGCAGGGAGCCCGGCTTGTGGGAAAAAAACCTTGTGGGGGTTCTGTG 5040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  5041 GCGCCTTGGACGGCGCTGAGCGCTGTCCGCTGGCCCAGAGCCCCAGGAGCCCAGATACCA 5100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041 GCGCCTTGGACGGCGCTGAGCGCTGTCCGCTGGCCCAGAGCCCCAGGAGCCCAGATACCA 5100

Qy  5101 AATGTGCTTCACTGAGTCAATCAATACTTCGTGGCTGTTGGCGTTGGAGT 5160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101 AATGTGCTTCACTGAGTCAATCAATACTTCTGGGACAGCCGCACTCGCTGTTGGCGTTGGAGT 5160

Qy  5161 GGCTATGCTTATCTAGCCATTGTGACACTTTGCGGCGCCCACTTTGTGTCGCGGTTCTGCGGTC 5220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161 GGCTATGCTTATCTAGCCATTGTGACACTTTGCGGCGCCCACTTTGTGTCGCGGTTCTGCGGTC 5220

Qy  5221 TATTACAATCAGTCCCTACCGTGTCCACCGTGTGACGTGGTGACGAAGAGAAATCGT 5280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221 TATTACAATCAGTCCCTACCGTGTCCACCGTGTGACGTGGTGACGAAGAGAAATCGT 5280

Qy  5281 GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGTGCTGCAATTGACAAGCTGAA 5340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281 GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGTGCTGCAATCGATAAGCTGAA 5340

Qy  5341 GAGTACAATCACCACAACTAGTGCTCTTTCACATTGGAAACCGCCTTGAAAACTTAACAC 5400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341 GAGTACAATAACCACAACTAGTCCTTTCACATTGGAAACCGCCTTGAAAAACTTAACAC 5400

Qy  5401 CTTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTTTAGT 5460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401 CTTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTTTAGT 5460

Qy  5461 CACTTTACCTGACAATCCCTTGTTCCTTCATTGGAGGGTATTACTAC 5520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461 CACTTTACCTGACAATCCCTTGTTCCTTCATTGGAGGGTATTACTAC 5520

Qy  5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTTGCCACTGACCGCCGCAATTGCGTCCAA 5580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTTGCCACTGACCGCCGCAATTGCGTCCAA 5580

Qy  5581 GCTTACAGACGCTAGAGGCGCACTGGCGTTCATGATGCCGCGGCTGCGGGAACAGTCT 5640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581 GCTTACAGACGCTAGAGGCGCACTGGCGTTCATGATGCCGCGGCTGCGGGAACAGTCT 5640

Qy  5641 TGGTACATGGACATCGGTGGGTTTGTCTTGACATGCTACATGCTCTAGGCGGCCTCGCCGCCTC 5700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641 TGGTACATGGACATCGGTGGGTTTGTCTTGACATGCTACATGCTCTAGGCGGCCTAGGCCGCCTC 5700

Qy  5701 ATCCACTGCTTGTTGACATTTAAATGCTTGAGTGAGTGAGATCAGCT 5760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701 ATCCACTGCTTGTTGACATTTAAATGCTTGAGTGAGTGAGATCAGCT 5760

Qy  5761 TGCTCGGTTAGTCTACTCCGCGGTTCAATCCGGCCGCAGGAGTTGTGGGCGTCTTGTCAGC 5820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761 TGCTCGGTTAGTCTACTCCGCGGTTCAATCCGGCCGCAGGAGTTGTGGGCGTCTTGTCAGC 5820

Qy  5821 TTGTGCAATGTTTGCTTTGACAACAGCAGGGGATGATTGCCCAACACTGCCCAACAGACTTCTTAC 5880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821 TTGTGCAATGTTTGCTTTGACAACAGCAGGGGATGATTGCCCAACACTGCCCAACAGACTTCTTAC 5880

Qy  5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGCCACTGTATTATTGCCACTCGTGACATCCG 5940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGCCACTGTATTATTGCCACTCGTGACATCCG 5940

Qy  5941 CAGGAAGATACTGGGCATTCTGGAGCATCTACCCCCTCGAGTGTCATATCAGCTTGCAT 6000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941 CAGGAAGATACTGGGCATTCTGGAGCATCTACCCCCTCGAGTGTCATATCAGCTTGCAT 6000

Qy  6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGGCCTCATTGCTTGGGGTCTAGAGAT 6060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGGCCTCATTGCTTGGGGTCTAGAGAT 6060

Qy  6061 TTGGCAGTATGTGCAATTTCTTTGCTTAATGTCTTTAAAGCTGATGAGTTCA 6120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  6061 TTGGCAGTATGTGTGCAATTTCTTGTGATTGCTTTAATGTCCTTAAAGCTGGAGTTCA   6120       Qy 7201 TTCCACTCAGTCAGCCCCGCAAACGGCCTACAAAAAGAAGAAGTGGGAAAGAGAGTGAGTT 7260
                                                                                         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6121 GAGCATGGTTAACATTCCTGGTTGTCGTCTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC 6180       Db 7201 TTCCACTCAGTCAGCCCCGCAAACGGCCTACAAAAAGAAGAAGTGGGAAAGAGAGTGAGTT 7260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121 GAGCATGGTTAACATTCCTGGTTGTCGTCTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC 6180       Qy 7261 TTCGTGAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT   7320
                                                                                         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6181 CTGGATTGGATCAGGTATGCTCCAAGCACGCGTGTCCATGCGGTGCTGAACTCATCTTTTC 6240       Db 7261 TTCGTGAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT   7320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181 CTGGATTGGATCAGGTATGCTCCAAGCACGCGTGTCCATGCGGTGCTGAACTCATCTTTTC 6240       Qy 7321 TCTGTCTGCAACTCGGGCCATCGGGCCATCACTAGTGTTCCTCAAACAAAGATCATTGGTGTATGT 7380
                                                                                         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6241 TGTTGAGAATGGTTTTGCAAAACTTTACAAAGGACCCAGAAACTTGTTCAAATTACTGGAG 6300       Db 7321 TCTGTCTGCAACTCGGGCCATCGGGCCATCACTAGTGTTCCTCAAACAAAGATCATTGGTGTATGT 7380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241 TGTTGAGAATGGTTTTGCAAAACTTTACAAAGGACCCAGAAACTTGTTCAAATTACTGGAG 6300       Qy 7381 GACTGAGCCGCGGCGGATGGGGAGCTTAGAAACAAAAGTCACTATTAATAGACAACCTCT  7440
                                                                                         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6301 AGGGGCTGTTCCAGTCAACGCTGTCTAGGCCGTGTGGGTTATCTAGGTTATCTAGGTGAGCCCAACTGATTG 6360   Db 7381 GACTGAGCCGCGGCGGATGGGGAGCTTAGAAACAAAAGTCACTATTAATAGACAACCTCT  7440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301 AGGGGCTGTTCCAGTCAACGCTGTCTAGGCCGTGTGGGTTATCTAGGTTATCTAGGTGAGCCCAACTGATTG 6360   Qy 7441 GTTCCCCATCATACCAAGCAAGCAAGTTGGCTAAGGAGAAAAGCTTCAAAAGTTGT     7500
                                                                                         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6361 GACTAGTCTTGTCGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAATGGGAGA 6420       Db 7441 GTTCCCCATCATACCAAGCAAGCAAGTTGGCTAAGGAGAAAAGCTTCAAAAGTTGT     7500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361 GACTAGTCTTGTCGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAATGGGAGA 6420       Qy 7501 CGGTGTCATGTGGGACTATGAAGTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTC   7560
                                                                                         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6421 TCACATTTTGTTACAGCAGTATCCTCTCCCAAATGTCTGTTTCACCCAGGTGCCCCCAAC   6480       Db 7501 CGGTGTCATGTGGGACTATGAAGTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTC   7560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421 TCACATTTTGTTACAGCAGTATCCTCTCCCAAATGTCTGTTTCACCCAGGTGCCCCCAAC   6480       Qy 7561 CCACATCACTGGCCTTCGGGGACTGATGTTCGTTCTTCGTTCTGAGCACGCCCAAGGCTGTTCT 7620
                                                                                         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6481 CTTGAGAGCTGCAGTGGCCGTGAACGGCGCGTACAGGTTCAGTGTTATCTAGGTAGGTGAGCCCAA 6540     Db 7561 CCACATCACTGGCCTTCGGGGACTGATGTTCGTTCTTCGTTCTGAGCACGCCCAAGGCTGTTCT 7620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481 CTTGAGAGCTGCAGTGGCCGTGAACGGCGCGTACAGGTTCAGTGTTATCTAGGTAGGTGAGCCCAA 6540     Qy 7621 GGACTTGCAGAAGTGTGTCGAGCGAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT  7680
                                                                                         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6541 AACTCCTTGACACACATCGCTTGCTTGCTCGACGTGTGTTGGTGTAAGGGTAAAACTGTTAA   6600       Db 7621 GGACTTGCAGAAGTGTGTCGAGCGAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT  7680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541 AACTCCTTGACACACATCGCTTGCTTGCTCGACGTGTGTTGGTGTAAGGGTAAAACTGTTAA   6600       Qy 7681 AGTTCCAAAGGAGGAGGTCTTCGTGTTGTCGTGTGAAGACCCCCAGAAACAACAAGAAACCCCCAAG 7740
                                                                                         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6601 GCTTCCCTTCCCGCGTTGACGGTCACACACCTGGTGTCGCGCATGCCAGTGCAACTTAATTGCGTGA 6660     Db 7681 AGTTCCAAAGGAGGAGGTCTTCGTGTTGTCGTGTGAAGACCCCCAGAAACAACAAGAAACCCCCAAG 7740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601 GCTTCCCTTCCCGCGTTGACGGTCACACACCTGGTGTCGCGCATGCCAGTGCAACTTAATTGCGTGA 6660     Qy 7741 GCTTATCTCGTACCCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGT  7800
                                                                                         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6661 TGGACTTGAGAGACAAATGACTGTAATTCCACACAACAACAACACTCCTAGTGATGAGCCGCAGT 6720     Db 7741 GCTTATCTCGTACCCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGT  7800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661 TGGACTTGAGAGACAAATGACTGTAATTCCACACAACAACAACACTCCTAGTGATGAGCCGCAGT 6720     Qy 7801 TGCCTCCTGACGTAGTTGTAAAGCTGTCATGGGAGATGCGTTGTGTCGACCCACGTAC   7860
                                                                                         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6721 GTCCGGCTCTGTTTGTTTCAAAAGTGTCATGTCAACGAAGTGACCGTTAGAAGTCCTTCAAACCT 6780     Db 7801 TGCCTCCTGACGTAGTTGTAAAGCTGTCATGGGAGATGCGTTGTGTCGACCCACGTAC   7860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721 GTCCGGCTCTGTTTGTTTCAAAAGTGTCATGTCAACGAAGTGACCGTTAGAAGTCCTTCAAACCT 6780     Qy 7861 CCGTGTCAAGCGTCGTCGTGTTCGTGATGGTGTCCGAGCCCATGCGCCACATGCGGATAC 7920
                                                                                         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6781 AGTCTCCTTCCACCACCTGTTCTACAGTTGGCCAGCCCCTCCATCGAGAGAGGTAGTGTAAGAAA 6840     Db 7861 CCGTGTCAAGCGTCGTCGTGTTCGTGATGGTGTCCGAGCCCATGCGCCACATGCGGATAC 7920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781 AGTCTCCTTCCACCACCTGTTCTACAGTTGGCCAGCCCCTCCATCGAGAGAGGTAGTGTAAGAAA 6840     Qy 7921 AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTGAGACAGACAGACATCTACTC 7980
                                                                                         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6841 GCGCCAGTCCGGGCAAGAACTTGGTTCGCTTGCTCAATGTGCAATGACCGAGAGCCGTCCC   6900       Db 7921 AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTGAGACAGACAGACATCTACTC 7980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841 GCGCCAGTCCGGGCAAGAACTTGGTTCGCTTGCTCAATGTGCAATGACCGAGAGCCGTCCC   6900       Qy 7981 AGCAGCTAAACTCAGTGACCGATGAAGAAGAACCTCGCTTCCTTATTGCGAGGGCAGTTATA 8040
                                                                                         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6901 AGGAGTGTCATGTCTACACCTGTTCACAGGTCTCAGACGGAGGTCAGACGAAAGTTGGTCGAC  6960      Db 7981 AGCAGCTAAACTCAGTGACCGATGAAGAAGAACCTCGCTTCCTTATTGCGAGGGCAGTTATA 8040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901 AGGAGTGTCATGTCTACACCTGTTCACAGGTCTCAGACGGAGGTCAGACGAAAGTTGGTCGAC  6960      Qy 8041 CGCTTGGAGGACCGATGATCGCTTATGATGGCCGAGAGACATCGGATATCGGATATCGGATATCGGATAC 8100
                                                                                         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  6961 CCCTCCTTCACCACCTGTTCTACAGTTGGCAATGCTTGGCCATGCCGATGCCGATGGGGAGCGGGTGA 7020    Db 8041 CGCTTGGAGGACCGATGATCGCTTATGATGGCCGAGAGACATCGGATATCGGATATCGGATATCGGATAC 8100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961 CCCTCCTTCACCACCTGTTCTACAGTTGGCAATGCTTGGCCATGCCGATGCCGATGGGGAGCGGGTGA 7020    Qy 8101 TTCCGGCGTCTATACTCAGTTCAAGTTCCAACAGTTGACCTGCTGAAGGTAAATGC    8160
                                                                                         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAAACAGGCGGAGGCCCTGATGA  7080        Db 8101 TTCCGGCGTCTATACTCAGTTCAAGTTCCAACAGTTGACCTGCTGAAGGTAAATGC    8160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAAACAGGCGGAGGCCCTGATGA  7080        Qy 8161 TGCAGCCGAACAGGCTGCGATGAAGAACCTCGCTTCCTTATTTGCGGCGGCGATGATTGCAC 8220
                                                                                         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  7081 TTTACCCAGTTACCCTCCCAAAAAGGAGTCTCGAACGTGGTCGACGTGGTCGAC        7140        Db 8161 TGCAGCCGAACAGGCTGCGATGAAGAACCTCGCTTCCTTATTTGCGGCGGCGATGATTGCAC 8220
         |||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  7081 TTTACCCAGTTACCCTCCCAAAAAGGAGTCTCGAACGTGGTCGACGTGGTCGAC        7140        Qy 8221 CGTAATTTGGAAGAGCGCGGAGATGCAGACAACAAGCAATGCGTGTCTTTGCTAG     8280
                                                                                         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  7141 GGCTACAACCGCTTCCAGCTACGTTACTGGCCCCCGTACCCCGTACCCCTAAGATACGGGAAAGGA 7200     Db 8221 CGTAATTTGGAAGAGCGCGGAGATGCAGACAACAAGCAATGCGTGTCTTTGCTAG     8280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141 GACTACAACCGCTTCCAGCTACGTTACTGGCCCCCGTACCCCGTACCCCTAAGATACGGGAAAGGA 7200
```

APPLICANT: JAMES C. ERKER
APPLICANT: SHERI L. BUIJK
APPLICANT: ISA K. MUSHAHWAR
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
NUMBER OF SEQUENCES: 716
CORRESPONDENCE ADDRESS:
 ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
 STREET: 100 ABBOTT PARK ROAD
 CITY: ABBOTT PARK
 STATE: IL
 COUNTRY: USA
 ZIP: 60064-3500
COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/488,446
 FILING DATE:
 CLASSIFICATION:
PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/424,550
 FILING DATE:
ATTORNEY/AGENT INFORMATION:
 NAME: POREMBSKI, PRISCILLA E.
 REGISTRATION NUMBER: 33,207
 REFERENCE/DOCKET NUMBER: 5527.PC.01
TELECOMMUNICATION INFORMATION:
 TELEPHONE: 708-937-6365
 TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 390:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9143 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
US-08-488-446-390

Query Match   96.4%;   Score 9059.8;   DB 4;   Length 9143;
Best Local Similarity 99.6%;   Pred. No. 0;
Matches 9096;   Conservative   0;   Mismatches   37;   Indels   4;   Gaps   1;

Qy   1 ACCACAAACACTTCCAGTTGTTACACTCCGTAGGAATGCTCCTGGAGCAGTGCTCCTGGAGCACCCCCCTAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCACAAACACTTCCAGTTGTTACACTCCGTAGGAATGCTCCTGGAGCAGTGCTCCTGGAGCACCCCCCTAG 60

Qy   61 CAGGGGCGTGGGGGATTCCCCTGCCCCGTCTGCAGAAGGGTGGAGCCAACCACCTTAGTAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CAGGGGCGTGGGGGATTCCCCTGCCCCGTCTGCAGAAGGGTGGAGCCAACCACCTTAGTAT 120

Qy   121 GTAGGCGGCGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTAGGCGGCGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC 180

Qy   181 CCTGATGGGCGTTCATGGGTTCGGTGGTGGTGGCGGCGCTTAGGCAGCCTCACGCGCCACCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCTGATGGGCGTTCATGGGTTCGGTGGTGGTGGCGGCGCTTAGGCAGCCTCACGCGCCACCA 240

Qy   241 CCTCCCAGATAGAGCGGCGGCACTGTAGGGAAGACCGGGAGACCGGGAGACCTCACTACCAAGGACG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCCCAGATAGAGCGGCGGCACTGTAGGGAAGACCGGGAGACCGGGAGACCTCACTACCAAGGACG 300

Qy   301 CAGACCTCTTTTTGAGTATCACGCTCCGGAGTAGTTGGGCAAGCCCACCTATATGTGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGACCTCTTTTTGAGTATCACGCTCCGGAGTAGTTGGGCAAGCCCACCTATATGTGT 360

Qy   361 TGGGATGGTTGGGGTTAGCCATTCCATATCCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TGGGATGGTTGGGGTTAGCCATTCCATATCCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT 420

Qy   8281 CTGGATGAAGGTGATGGGTGACCACCAGAGATTGTGTGCCTCAACCCAAAATACAGTTTGGA 8340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8281 CTGGATGAAGGTGATGGGTGACCACCAGAGATTGTGTGCCTCAACCCAAAATACAGTTTGGA 8340

Qy   8341 AGAATAACATCATCAATGTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8341 AGAATAACATCATCAATGTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400

Qy   8401 CTACTTTCTTACAAGAGATCTCGTATCCCCCTTGGCAGGTGCTCTGCCGAGGGTCTCGGG 8460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTGCCGAGGGTCTCGGG 8460

Qy   8461 ATACAACCCCAGTGCGTGCGTGGATTGGGTATCTAATACATCACTACCCATGTTTGTGGGT 8520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8461 ATACAACCCCCAGTGCGTGCGTGGATTGGGTATCTAATACATCACTACCCATGTTTGTGGGT 8520

Qy   8521 TAGCCGTGTGTGTGGCCTGTCAGTTCCATTTCATGGAGCAGATGCTCTTGAGGACAAACTTCCCGA 8580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8521 TAGCCGTGTGTGTGGCCTGTCAGTTCCATTTCATGGAGCAGATGCTCTTGAGGACAAACTTCCCGA 8580

Qy   8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640

Qy   8641 CATCATTGCTGCTGTGTGTGCACGGTATTGAGGCTTTCTCCGGTGCGCTACACCAACGCTGA 8700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8641 CATCATTGCTGCTGTGTGTGCACGGTATTGAGGCTTTCTCCGGTGCGCTACACCAACGCTGA 8700

Qy   8701 GATCCTCAGAGTTCCAATCACTAACAGACATGACCATGCCCCCCTGCCGAGCCTTGGCG 8760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8701 GATCCTCAGAGTTCCAATCACTAACAGACATGACCATGCCCCCCTGCCGAGCCTTGGCG 8760

Qy   8761 AAAGAAAGCCAGGGCCGGTCCTCTGCCCAGCGCCAAGAGGCGTGGCGGAGCACACGCAAAATT 8820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8761 AAAGAAAGCCAGGGCCGGTCCTCTGCCCAGCGCCAAGAGGCGTGGCGGAGCACACGCAAAATT 8820

Qy   8821 GGCTCGCTTCCTTCCTTCTCGGCATGCTACATCTAGACCTCTACCAGATTGGATAAGACGAG 8880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8821 GGCTCGCTTCCTTCCTTCTCGGCATGCTACATCTAGACCTCTACCAGATTGGATAAGACGAG 8880

Qy   8881 CGTGGCTCGGTACACCACTTCAATTATTGTGATGTTTACTCCCCGAGGGGGATGTGTT 8940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8881 CGTGGCTCGGTACACCACTTCAATTATTGTGATGTTTACTCCCCGAGGGGGATGTGTT 8940

Qy   8941 TATTACCACCAGAGAGATTGCAGAGAGTTCCTTGTGAAGTATTGGCCTGTCATTGTTTT 9000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8941 TGTTACCACCAGAGAGATTGCAGAGAGTTCCTTGTGAAGTATTGGCCTGTCATTGTTTT 9000

Qy   9001 TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA 9060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9001 TGCCCTAGGGCTCATTGCTGTTGGACTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA 9060

Qy   9061 CTAACAG---TTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC 9116
          ||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC 9120

Qy   9117 GGGCTTAACGACCCCGC 9133
          |||||||||||||||||
Db   9121 GGGCTTAACGACCCCGC 9137

RESULT 4
US-08-488-446-390
; Sequence 390, Application US/08488446
; Patent No. 6558898
GENERAL INFORMATION:
 APPLICANT: JOHN N. SIMONS
 APPLICANT: TAMI J. PILOT-MATIAS
 APPLICANT: GEORGE J. DAWSON
 APPLICANT: GEORGE G. SCHLAUDER
 APPLICANT: SURESH M. DESAI
 APPLICANT: THOMAS P. LEARY
 APPLICANT: ANTHONY SCOTT MUERHOFF

```
Qy   421  CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTTCTACTCAAACAAGTTCCTGTACC  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTTCTACTCAAACAAGTTCCTGTACC  480

Qy   481  TGCGCCCAGAACGCGCAAGAACAAGCAGCAGGCTTCATATCCTGTGTCCATTAAAAC      540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  TGCGCCCAGAACGCGCAAGAACAAGCAGCAGGCTTCATATCCTGTGTCCATTAAAAC      540

Qy   541  ATCTGTTGAAAGGGGACAACAGCAGCAAAGTCCAGCGGATGCTCGGCCTCGTAA          600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  ATCTGTTGAAAGGGGACAACAGCAGCAAAGTCCAGCGGATGCTCGGCCTCGTAA          600

Qy   601  TTACAAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTGCCAGC    660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  TTACAAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTGCCAGC    660

Qy   661  TCATGGTTGGGACGGCGTCGTTCGACCAGTCTGCCATAAGTCTTCGCAATCTTCTGGA      720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  TCATGGTTGGGACGGCGTCGTTCGACCAGTCTGCCATAAGTCTTCGCAATCTTCTGGA      720

Qy   721  TTACCCTTGGGGTGGATTGGTGATGTTACAACTCACACACCTTCTAGTAGGCCCGCTGGT    780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  TTACCCTTGGGGTGGATTGGTGATGTTACAACTCACACACCTTCTAGTAGGCCCGCTGGT    780

Qy   781  GGCAGGAGGCGGTCGTTCGACCAGTCTGCCAGATAGTACGCTTGCTGGAGGATGGAGTCAA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  GGCAGGAGGCGGTCGTTCGACCAGTCTGCCAGATAGTACGCTTGCTGGAGGATGGAGTCAA  840

Qy   841  CTTGGGCTACTGGTTGGTTCGGTGTCCCACCTTTTTGTGGTATGTCTGCTATCTTTGGGCTG  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  CTTGGGCTACTGGTTGGTTCGGTGTCCCACCTTTTTGTGGTATGTCTGCTATCTTTGGGCTG  900

Qy   901  TCCCTGTAGTAGTGGGGCGCGGGTCACTGACCCCAGACACAAATACCACCAATTG          960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  TCCCTGTAGTAGTGGGGCGCGGGTCACTGACCCCAGACACAAATACCACCAATTG          960

Qy   961  CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCTTGCCACTGCTACACGAGCCTGGTTG    1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCTTGCCACTGCTACACGAGCCTGGTTG    1020

Qy  1021  TGTGATCTGTGCCGACGAGTGCTGTGGGTTTCCCGCCAATCGTACATCTTCACACCTTCCAA  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  TGTGATCTGTGCCGACGAGTGCTGTGGGTTTCCCGCCAATCGTACATCTTCACACCTTCCAA  1080

Qy  1081  TTGGACTGGCACCGGACTCCTTCTTCTGGCTGACCACCATTGATTTTGTTATGGGGCTCTGT  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  TTGGACTGGCACCGGACTCCTTCTTCTGGCTGACCACCATTGATTTTGTTATGGGGCTCTGT  1140

Qy  1141  GACCTGTGACGCGCCCTTGACATTGGTGAGTTCGGTGTGTATTAGTCGGTGACTG        1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  GACCTGTGACGCGCCCTTGACATTGGTGAGTTCGGTGTGTATTAGTCGGTGACTG        1200

Qy  1201  GCTTGTGTCAGGCACGGACTTATTCACAATAGACCTCAATGAAACTGGTACTTGTTGTACCTGGA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GCTTGTGTCAGGCACGGACTTATTCACAATAGACCTCAATGAAACTGGTACTTGTTGTACCTGGA  1260

Qy  1261  AGTGCCCACTGGAATAGATCCTGGGGTTATCGGGTTATGCGGTGGATGGCCGGCAAGGT    1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  AGTGCCCACTGGAATAGATCCTGGGGTTATCGGGTTATGCGGTGGATGGCCGGCAAGGT    1320

Qy  1321  CGAGGCTGTCATCTTCTTGACCAAACTGCTTCACAAGTGGCTTACCATACGCTATTGCGACTAT  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  CGAGGCTGTCATCTTCTTGACCAAACTGCTTCACAAGTGGCTTACCATACGCTATTGCGACTAT  1380

Qy  1381  GTTTAGCAGTGTACACTACCTGCGGGTTGGCGCTTCTGATCTACTAGCTTCGCGGGGCAA    1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  GTTTAGCAGTGTACACTACCTGCGGGTTGGCGCTTCTGATCTACTAGCTTCGCGGGGCAA    1440

Qy  1441  GTGGTATCAGTTGCTCCTCTAGCGCTTATGCTTTACATGAAGCGACCTTCTGGAAACCCCAT  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  GTGGTATCAGTTGCTCCTCTAGCGCTTATGCTTTACATGAAGCGACCTTCTGGAAACCCTAT  1500

Qy  1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTTGTCGCCCTTTGATGATACCATGTCC
```

```
Db  1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTTGTCGCCCTTTGATGATACCATGTCC  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1561  TTGCCACTCTTATTTGAGTGAGAATGTCAGAGAATCATTTGTTACAGTCCAAAGTGGAC    1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  TTGCCACTCTTATTTGAGTGAGAATGTCAGAGAATCATTTGTTACAGTCCAAAGTGGAC    1620

Qy  1621  CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG   1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  CAGGCCTGTCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG   1680

Qy  1681  TGGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTGCTGCCGTATTCGCAA    1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  TGGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTGCTGCCGTATTCGCAA    1740

Qy  1741  TGTGCCATCGTACTGCACTATGGGCACTCAGTGTGGAACGACACTCGCAACACTTA       1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  TGTGCCATCGTACTGCACTATGGGCACTCAGTGTGGAACGACACTCGCAACACTTA       1800

Qy  1801  CGAAGCATGCGGTGTAACACCATGGCTAACAACCGCACATGGCACAACGGCTCAGCCTGAA  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  CGAAGCATGCGGTGTAACACCATGGCTAACAACCGCACATGGCACAACGGCTCAGCCTGAA  1860

Qy  1861  ATTGGCTATATTACAATACCCTGGGTCTAAAGAAAATGTTAAACCTCATAATTGGATGTC    1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  ATTGGCTATATTACAATACCCTGGGTCTAAAGAAAATGTTAAACCTCATAATTGGATGTC    1920

Qy  1921  AGGCCATTGTATTTGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAA     1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  AGGCCATTGTATTTGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAA     1980

Qy  1981  TTTCACTCTCCTACCACCGGAGAGGTGGCTAGGTTGCCCGGTACCCACCTGTGGTACG    2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  TTTCACTCTCCTACCACCGGAGAGGTGGCTAGGTTGCCCGGTACCCACCTGTGGTACG    2040

Qy  2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAGG    2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAGG    2100

Qy  2101  ATTGATCACCAAAGACAAAGCCTGGAACAAGCCGTGGTGTACTTCAGGTCTTATATTCAA   2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  ATTGATCACCAAAGACAAAGCCTGGAACAAGCCGTGGTGTACTTCAGGTCTTATATTCAA   2160

Qy  2161  TTTGTCTCTTACGGGAGTTACCACCAAGCCGTTGGTCGTTGTTCTGTGGGGTTGTGTGG    2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TTTGTCTCTTACGGGAGTTACCACCAAGCCGTTGGTCGTTGTTCTGTGGGGTTGTGTGG    2220

Qy  2221  CAGCAAGTATCTTATTTTAGCCTTTTTAGCCTTTGTTTTGGGGCCGCTTC            2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  CAGCAAGTATCTTATTTTAGCCTTTTTAGCCTTTGTTTTGGGGCCGCTTC            2280

Qy  2281  TGGTTACCCTTTGGCTCCTTGCTGTCTCCATCCGTATCTCGAATCTCGTAATTCTCTTGT   2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TGGTTACCCTTTGGCTCCTTGCTGTCTCCATCCGTATCTCGAATCTCGTAATTCTCTTGT   2340

Qy  2341  TTTGTCTAAAGCTCAAGTAGCTCAAGTAGCTCCTTTTGATTTTTCATCTGTTGCTATCTCCG  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  TTTGTCTAAAGCTCAAGTAGCTCAAGTAGCTCCTTTTGATTTTTCATCTGTTGCTATCTCCG  2400

Qy  2401  CTGCAGGCCACGTTATGCCGCTGCCCCATGGTTGTGCCCATGCTGCGCCCCCT         2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  CTGCAGGCCACGTTATGCCGCTGCCCCATGGTTGTGCCCATGCTGCGCCCCCT         2460

Qy  2461  AACTTTCTTTGTTGCAGCAGCTGCTCGTGGGCCAACCAGATTATGACTGCTGTGGGACTGCT  2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  AACTTTCTTTGTTGCAGCAGCTGCTCGTGGGCCAACCAGATTATGACTGCTGTGGGACTGCT  2520

Qy  2521  AGTGGCAGGGTAGTTTTGTGGGCCGGCCGTAACCGTGGTGCTGGTGGTCATAGCTCTTGT  2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  AGTGGCAGGGTAGTTTTGTGGGCCGGCCGTAACCGTGGTGCTGGTGGTCATAGCTCTTGT  2580

Qy  2581  AGTGCCTTGGCCTCGATGCTCAATAGCTGAGTTTTGTCGCGGCCTTTAACCCTCTTGGTTACGCTGTTCAGC  2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy  3721 GATGTTCACCGCTGCTAGAAATTCTGGCGGTTCAGTCAGTCAGATTCAGTTAGGCCGTT 3780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721 GATGTTCACCGCTGCTAGAAATTCTGGCGGTTCAGTCAGTCAGATTCAGTTAGGCCGTT 3780

Qy  3781 GGTGTGTGCTGGATACCATCCCAGTACACAGACATGCCACTCTTGATACAAAACCTAC 3840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781 GGTGTGTGCTGGATACCATCCCAGTACACAGACACATGCCACTCTTGATACAAAACCTAC 3840

Qy  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCACTGGCAGCGGCAAGTCAAC 3900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCACTGGCAGCGGCAAGTCAAC 3900

Qy  3901 CAAATTACCACTTCTTCACATGCACGAGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT 3960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 CAAATTACCACTTCTTCACATGCACGAGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT 3960

Qy  3961 GGCTACAACAGCATCAATGCCAAAGTACAACACAGGGGCTTCACTTAGTACACGTACAGTGGCAT 4020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 GGCTACAACAGCATCAATGCCAAAGTACAACACAGGGGCTTCACTTAGTACACGTACATATGGCAT 4020

Qy  4021 CTATTTTAATGCAAATGTACCAACACACAGGGGCTTCACTTGTGACGAATGCCATGC 4080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 CTATTTTAATGCAAATGTACCAACACACAGGGGCTTCACTTGTGACGAATGCCATGC 4080

Qy  4081 GTACCTGACCGGACATGTTCCCGGACATGTGAGTCATGATGTAATCATTGTGACGACACTTGCA 4140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 GTACCTGACCGGACATGTTCCCGGACATGTGAGTCATGATGTAATCATTGTGACGACACTTGCA 4140

Qy  4141 TACCGATGCAACCACCGTGTTCTTGCCACGGCTTGGCGTGTACCCCCCATCCCAAAAA 4200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 TACCGATGCAACCACCGTGTTCTTGCCACGGCTTGGCGTGTACCCCCCATCCCAAAAA 4200

Qy  4201 TGTTAGGCTAGTGGTCTTGCGCACGGCTACCCCTGGAGTAATCCCTACCCACCACATGC 4260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201 TGTTAGGCTAGTGGTCTTGCGCACGGCTACCCCTGGAGTAATCCCTACCCACCACATGC 4260

Qy  4261 CAACATAACTACTGAGATTCAATTAACCGATGAAGGAGCACTCATCCCCTTTCATGTGGAAAAAGAT 4320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261 CAACATAACTACTGAGATTCAATTAACCGATGAAGGAGCACTCATCCCCTTTCATGTGGAAAAAGAT 4320

Qy  4321 TAAGGAGGAAAATCTGAAGAAAGGGAGACACCTTATCTCGAGGCTACCAAAAAACACTG 4380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321 TAAGGAGGAAAATCTGAAGAAAGGGAGACACCTTATCTCGAGGCTACCAAAAAACACTG 4380

Qy  4381 TGATGAGCTTGCTAACGAGTTAGCTGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381 TGATGAGCTTGCTAACGAGTTAGCTGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440

Qy  4441 ATGTGACATCTCAAAAATCCCGAGGCGACTGTGAGTAGTGCCACTGATGCCTTGTG 4500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441 AATGTGACATCTCAAAAATCCCGAGGCGACTGTGAGTAGTGCCACTGATGCCTTGTG 4500

Qy  4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGAGCTCATGGTGCAGCCTCATGGTAGAAGGCAC 4560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGAGCTCATGGTGCAGCCTCATGGTAGAAGGCAC 4560

Qy  4561 ATGCCATGTTGACCTTGACCCCTACTTTCACCATGGGGTTCGTGTTCGTGTGCGGGGTTTCAGC 4620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561 ATGCCATGTTGACCTTGACCCCTACTTTCACCATGGGGTTCGTGTTCGTGTGCGGGGTCTCAGC 4620

Qy  4621 AATAGTTAAAGGCCAGCGGATATCGGGAGAGCTGGCATCATATACTACTA 4680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621 AATAGTTAAAGGCCAGCGGATATCGGGAGAGCTGGCATCATATACTACTA 4680

Qy  4681 TGTAGACGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681 TGTAGACGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740

Qy  4741 CGACGCAGCCAAGGCATGGTATGGTTTGTCATCAACAGAAGCTCAAACTATTCTGGACAC 4800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741 CGACGCAGCCAAGGCATGGTATGGTTTGTCATCAACAAGAAGCTCAAACTATTCTGGACAC 4800
```

```
Db  2581 AGGTCCTTGGCCTCTGGTAGCGCTTTTAACCCTCTTGCATTGGCTACGCCTGCTCAGC 2640

Qy  2641 TTTTGATACCGGAGATAATTGGAGGGCTGACATACACCCTGTAGTAGCATTAGTTGTCAT 2700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641 TTTTGACACCGGAGATAATTGGAGGGCTGACATACACCCTGTAGTAGCATTAGTTGTCAT 2700

Qy  2701 GTCTCGTTTGGCTTCTTGCTCACTTGTTACCTCGCTCGTTGCTTTAGTTAACTCCTATCT 2760
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 GTCTCGTTTGGCTTCTTGCTCACTTGTTACCTCGCTCGTTGCTTTAGTTAACTCCTATCT 2760

Qy  2761 TTGGCAACGTGGGAGAATTGGTTTGGAACGTTACACTAAGACCGGAGAGGTTTTTCCT 2820
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 TTGGCAACGTGGGAGAATTGGTTTGGAACGTTACACTAAGACCGGAGAGGTTTTTCCT 2820

Qy  2821 TGTGCTGTTGTTTCCCGGTGCGACATGAGCACGCGCTGGTGACTTCTGTCGTGTGTCA 2880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821 TGTGCTGTTGTTTCCCGGTGCGACATGAGCACGCGCTGGTGACTTCTGTCGTGTGTCA 2880

Qy  2881 CGTAGCTCTCTATGTTAACATCCAGTGACATCCAGCATCGTCTTCTTGGACTGACTCTAGGGT 2940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881 CGTAGCTCTCTATGTTAACATCCAGTGACATCCAGCATCGTCTTCTTGGACTGACTCTAGGGT 2940

Qy  2941 TAGGGCCCATAGAATGTTGGTCGTCTCGGAAAGTGTCATGCTTGACTTTGTGTGTTCATTATGT 3000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 TAGGGCCCATAGAATGTTGGTCGTCTCGGAAAGTGTCATGCTTGACTTTGTATTCTCATTATGT 3000

Qy  3001 TCTTAAGTTTTTCCCTCTAGTGTTTGGTGAGAATGGTGTGTTTTTCTATAAGCACTTGCA 3060
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 TCTTAAGTTTTTCCCTCTAGTGTTTGGTGAGAATGGTGTGTTTTTCTATAAGCACTTGCA 3060

Qy  3061 TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCAAGAGCCATTTTCCC 3120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCAAGAGCCATTTTCCC 3120

Qy  3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAGAGACGCTTGGCGTGTGGGACAC 3180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAGAGACGCTTGGCGTGTGGGACAC 3180

Qy  3181 GGTTGATGGTTTGCCCGTTGTTGCGCGCTCTGTGGTTTTCGCAGGGTTGGCTAT 3240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 GGTTGATGGTTTGCCCGTTGTTGCGCGCTCTGTGGTTTTCGCAGGGTTTAGCTAT 3240

Qy  3241 GCCGCCAGATGGGTGGCCATTACCGCACCTTTACGCTGCGAGTGTCTCTGAACGTGG 3300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 GCCGCCAGATGGGTGGCCATTACCGCACCTTTACGCTGCGAGTGTCTCTGAACGTGG 3300

Qy  3301 CACGCTGTCAGCGATGCGATGGTCATGACTGTATAGACCCCGAACTTGGACTGGAAC 3360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 CACGCTGTCAGCGATGCGATGGTCATGACTGTATAGACCCCGAACTTGGACTGGAAC 3360

Qy  3361 TATCTTCAGATTAGGAGATCTCTGGCCACTAGCTACACATGGATTGTTGTGTGACAAACGTGTT 3420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 TATCTTCAGATTAGGAGATCTCTGGCCACTAGCTACACATGGATTGTTGTGTGACAAACGTGTT 3420

Qy  3421 GTATACTGCTCACCATGGCAGCAAGGGGCGCCGGTTGGCTCATCCCACAGGCTCTATACA 3480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421 GTATACTGCTCACCATGGCAGCAAGGGGCGCCGGTTGGCTCATCCCACAGGCTCCATACA 3480

Qy  3481 CCCAATAACCGTTGACGCGGTAATGACCAGGATACATCTATCAACCACCATGTGGAGCTGG 3540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481 CCCAATAACCGTTGACGCGGTAATGACCAGGATACATCTATCAACCACCATGTGGAGCTGG 3540

Qy  3541 GTCCCTTACTCGGTCTCGTGCTCTTGCGGGAGACCAAGGGGTATCTGGTAACACGACTGGGGTC 3600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541 GTCCCTTACTCGGTCTCGTGCTCTTGCGGGAGACCAAGGGGTATCTGGTAACACGACTGGGGTC 3600

Qy  3601 ATTGGTTGAGGTCAACAAATCCGATGCCCATCAGGTGCGTGTGTGTGTCCTCGGGCATGTTATTGG 3660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601 ATTGGTTGAGGTCAACAAATCCGATGCCCATCAGGTGCGTGTGTGTGTCCTCGGGCATGTTATTGG 3660

Qy  3661 GGCTGTGTGCCAAGGGTTCTTCAGGTGCCTCCCCGGGCATGTTATTGG 3720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661 GGCTGTGTGCCAAGGGTTCTTCAGGTGCCTCCCCGGGCATGTTATTGG 3720
```

```
Qy   4801  CTATCGCACCCAACCTGGGTTACCTGCCGATAGGAGCAAATTTGACGAGTGGGCTGATCT  4860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801  CTATCGCACCCAACCTGGGTTACCTGCCGATAGGAGCAAATTTGACGAGTGGGCTGATCT  4860

Qy   4861  CTTTTCTATGGTCAACCCGAACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA    4920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861  CTTTTCTATGGTCAACCCGAACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA    4920

Qy   4921  TTATGTTTGTTGACTGCAGCAGCCAACTACAACTCTGTCATCAGTATGCGTATGCTGCTCC  4980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921  TTATGTTTGTTGACTGCAGCAGCCAACTACAACTCTGTCATCAGTATGCGTATGCTGCTCC  4980

Qy   4981  CAATGACCCACCACGGTGCAGGGAGCCCGGCTTGGGAAAAAACCTTGTGTGGGGTTCTGTG  5040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981  CAATGACCCACCACGGTGCAGGGAGCCCGGCTTGGGAAAAAACCTTGTGTGGGGTTCTGTG  5040

Qy   5041  GCGCTTGACGCGGCGCTGACGCCCTGCCTCCTGGCCCAGAGCCCAGCGAGGTGACCAGATACCA  5100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041  GCCCTATGCTTGACGCGGCGCTGACGCCCTGCCTCCTGGCCCAGAGCCCAGCGAGGTGACCAGATACCA  5100

Qy   5101  AATGTGCTTCACTGAAGTCAATATCTTCTGGGACACGCCCACTCGCGTTGGCGTTGGGAGT  5160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101  AATGTGCTTCACTGAAGTCAATATCTTCTGGGACACGCCCACTCGCGTTGGCGTTGGGAGT  5160

Qy   5161  GGCTATGGCTTATCTGACACTTTTGACACTTTTGGCGCGCCACTCGTGTGTGCGCGGTTGCTCGGTC  5220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161  GGCTATGGCTTATCTGACACTTTTGACACTTTTGGCGCGCCACTCGTGTGTGCGCGGTTGCTCGGTC  5220

Qy   5221  TATTACATCAGTCCCTACGGGTGCTACTCCTTGCTATCATAGAGTATTGCTGTGGTTGTGTCGTGAAATCGT  5280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221  TATTACATCAGTCCCTACGGGTGCTACTCCTTGCTATCATAGAGTATTGCTGTGGTTGTGTCGTGAAATCGT  5280

Qy   5281  GGAGGAGTGCATCATTCCCTTGCATCATGCGTGTTTGCTTTCATTGGCGTGCAATTGACAGCTGAA  5340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5281  GGAGGAGTGCATCATTCCCTTGCATCATGCGTGTTTGCTTTCATTGGCGTGCAATTGACAGCTGAA  5340

Qy   5341  GAGTACAATCACCACAGTCCTTCCTTCACATTGGAAACGCCCTTGAAAAACTTAACAC     5400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5341  GAGTACAATCACCACAGTCCTTCCTTCACATTGGAAACGCCCTTGAAAAACTTAACAC     5400

Qy   5401  CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTCGTGTGTTAGT       5460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5401  CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTCGTGTGTTAGT       5460

Qy   5461  CACTTTACCTGACAATCCCTTTGCATCATGACTGTGTTTGCTTTCATTGCGGGGTATTACTAC   5520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5461  CACTTTACCTGACAATCCCTTTGCATCATGACTGTGTTTGCTTTCATTGCGGGGTATTACTAC   5520

Qy   5521  CCCACTACCTACAAGATCAAAATGTTCCTGTCATTATTTGGAGGCGGCAATTGCGTCCAA     5580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5521  CCCACTACCTACAAGATCAAAATGTTCCTGTCATTATTTGGAGGCGGCAATTGCGTCCAA     5580

Qy   5581  GCTTACAGACGCTAGAGGCCACTGGTCGTTCATGATGGCCGGCGGCTGCGCGGAACAGCTCT    5640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5581  GCTTACAGACGCTAGAGGCCACTGGTCGTTCATGATGGCCGGCGGCTGCGCGGAACAGCTCT    5640

Qy   5641  TGGTGCAATGTTTGCTTTGACAACAGCAGGCCAGATCACTCGGCCAACACTACTTCTTAC    5700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5641  TGGTGCAATGTTTGCTTTGACAACAGCAGGCCAGATCACTCGGCCAACACTACTTCTTAC    5700

Qy   5701  ATCCACTGCTTGCACAGATCAAAATGCTTAAATGCTTGATGGGTGAGTGAGCCACTATGCAGCT  5760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5701  ATCCACTGCTTGCACAGATCAAAATGCTTAAATGCTTGATGGGTGAGTGAGCCACTATGCAGCT  5760

Qy   5761  TGCTGGTTTAGTCTACTCCGCGTTCAATCCGCGTCTTGTGCCGCGAGTTGTGTCAGC      5820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5761  TGCTGGTTTAGTCTACTCCGCGTTCAATCCGCGTCTTGTGCCGCGAGTTGTGTCAGC      5820

Qy   5821  TTGTGCAATGTTTGCTTTGACAACAGCAGGCCAGATCACTCGGCCAACACTTCTTAC      5880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5821  TTGTGCAATGTTTGCTTTGACAACAGCAGGCCAGATCACTCGGCCAACACTTCTTAC      5880

Qy   5881  TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATTGCAT  5940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5881  TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATTGCAT  5940

Qy   5941  CAGGAAGATACTGGCATTCTGGAGGCATCTACCCCTGGAGTGTCATATCAGCTTGCAT     6000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5941  CAGGAAGATACTGGCATTCTGGAGGCATCTACCCCTGGAGTGTCATATCAGCTTGCAT     6000

Qy   6001  CCGTTGGCTCCACACCCGACGGAGGAGGATGATTGCGGGCCTCATTGCTTGGGGTCTCAGAGAT  6060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6001  CCGTTGGCTCCACACCCGACGGAGGAGGATGATTGCGGGCCTCATTGCTTGGGGTCTCAGAGAT  6060

Qy   6061  TTGGCAGTATGTGTGCAATTTCTTGTGATTTGTATTGCTTAATGTCCTTAAAGCTGGAGTTCA  6120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6061  TTGGCAGTATGTGTGCAATTTCTTGTGATTTGTATTGCTTAATGTCCTTAAAGCTGGAGTTCA  6120

Qy   6121  GAGCATGGTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC     6180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6121  GAGCATGGTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC     6180

Qy   6181  CTGATTGGATCAGTAGGTGCTCCAAGCACGCTGTCCATGCGGTGCTGACAACTCATCTTTC  6240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6181  CTGATTGGATCAGTAGGTGCTCCAAGCACGCTGTCCATGCGGTGCTGACAACTCATCTTTC  6240

Qy   6241  TGTTGAGAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTACTGGAG  6300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6241  TGTTGAGAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTACTGGAG  6300

Qy   6301  AGGGGCTGTTCCAGTCAACGCTAGGCTGTGTGGGTCGGCTAGACCGCTGAACCTGATTG    6360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6301  AGGGGCTGTTCCAGTCAACGCTAGGCTGTGTGGGTCGGCTAGACCGCTGAACCTGATTG    6360

Qy   6361  GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA   6420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6361  GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA   6420

Qy   6421  TCACATTTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTGTTTCACCCAGTGCCCCAAC   6480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6421  TCACATTTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTGTTTCACCCAGTGCCCCAAC   6480

Qy   6481  CTTGAGAGCTGCAGTGGCCGTGACGGCCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA   6540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6481  CTTGAGAGCTGCAGTGGCCGTGACGGCCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA   6540

Qy   6541  AACTCCTTGGACGACATCTGCTTGCTGTTACGGTTACGGTAAAGGTAAAACTGTTAA      6600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6541  AACTCCTTGGACGACATCTGCTTGCTGTTACGGTTACGGTAAAGGTAAAACTGTTAA      6600

Qy   6601  GCTTCCCTTCCGCGGTTGACGGTGCGCATGCCAGCCCCGTGCGCGTGTGCCATGCGCGTGACACCTGGTGA  6660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6601  GCTTCCCTTCCGCGGTTGACGGTGCGCATGCCAGCCCCGTGCGCGTGTGCCATGCGCGTGACACCTGGTGA  6660

Qy   6661  TGCACTTGAGACAAATGACTGTAATTCCACAAACAACACTCCAGTGATGATGAAGGCGCAGT  6720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6661  TGCACTTGAGACAAATGACTGTAATTCCATAAACAACACTCCAGTGATGATGAAGGCGCAGT  6720

Qy   6721  GTCCGCTCTTGTTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC   6780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6721  GTCCCGCTCTTGTTTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC   6780

Qy   6781  AGCTGGCGGTTGACACCAAAACTGCCAGCCCCATCCAATCGAAGAGGTAGTGGTAAGAAA   6840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6781  AGCTGGCGGTTGACACCAAAACTGCCAGCCCCCATCCAATCGAAGAGGTAGTGGTAAGAAA   6840

Qy   6841  GCGCCAGTTCCGGCAAGAACTGTTCGCTTACTTGCCTTCCTCCCCTTAACCTGCGTCCC   6900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6841  GCGCCAGTTCCGGCAAGAACTGTTCGCTTACTTGCCTTCCTCCCCCTTACCTGCGTCCC   6900

Qy   6901  AGGAGTGTCATGTCCGTGAAAGCTCGAAAGTGAACGAAGTGCTGGAGTCCTTCAAAACCT  6960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6901  AGGAGTGTCATGTCTCCGTGAAAGCTCGAAAGTGCAACGAAGTGCTGACCCGTTAGAAAGTCCTTCAAACCT  6960

Qy   6961  CCCTCCTTCCACCACCTGTTCTACAGTTGGGCATGCCGATGCCCCTTGTTTGGGGAGCCGGGTGA  7020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6961  CCCTCCTTCCACCACCTGTTCTACAGTTGGGCATGCCGATGCCCCTTGTTTGGGGAGCCGGGTGA  7020
```

```
Db  CCCTTCTTCACCACCTGTTCTACAGTTGGCCATGCCGATGCCCCTGTTGGGAGCAGGTGA  7020
Qy  GTGTAACCCTTTCACTGCAATTGGATGTGCAATGCGAAACAGGCGGAGGCCCTGATGA    7080
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  GTGTAACCCTTTCACTGCAATTGGATGTGCAATGCGAAACAGGCGGAGGCCCTGATGA    7080
Qy  TTTACCCAGTTACCCTCCCAAAAAGAGAGGTCTCTGAATGTCAGACGAAGTTGGTCGAC  7140
    |||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db  TTTACCCAGTTACCCTCCCAAAAAGAGAGGTCTCTGAATGGTCAGACGAAGTTGGTCAAC  7140
Qy  GGCTACAACCGCTTCAGCTACGTTACTGGCCCCCCGTACCCTAAGATACGGGGAAAGGA  7200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  GACTACAACCGCTTCAGCTACGTTACTGGCCCCCCGTACCCTAAGATACGGGGAAAGGA  7200
Qy  TTTCCACTCAGTCAGCCCCGCCAAACGGCTACAACGGCTTACAGAAGAGTTGGAAAGAGTGAGTT  7260
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TTTCCACTCAGTCAGCCCCGCCAAACGGCTACAACGGCTTACAGAAGAGTTGGAAAGAGTGAGTT  7260
Qy  TTCGTGCAGCATGAGCTGACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCAAAGT  7320
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TTCGTGCAGCATGAGCTGACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCAAAGT  7320
Qy  TCTGTCTGCAACTCGGGCCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGT  7380
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TCTGTCTGCAACTCGGGCCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGT  7380
Qy  GACTGAGCCGCGGGATGCGGAGCTTAGAAACAAAAAGTCACTATTAATAGACAACCTCT  7440
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  GACTGAGCCGCGGGATGCGGAGCTTAGAAACAAAAAGTCACTATTAATAGACAACCTCT  7440
Qy  GTTCCCCCATCATACCACAAGCAAGCAAGTGAGATTGGCTAAGGAAAAGCTTCAAAGTTGT  7500
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  GTTCCCCCATCATACCACAAGCAAGCAAGTGAGATTGGCTAAGGAAAAGCTTCAAAGTTGT  7500
Qy  CGGTGTCATGTGGGACTATGATGAAGTAGCAGCTCACACGGTCGTGTAAGTCTGCTAAGTC  7560
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CGGTGTCATGTGGGACTATGATGAAGTAGCAGCTCACACGGTCGTGTAAGTCTGCTAAGTC  7560
Qy  CCACATCACTGCCCTTCGGGGCACTGATGTTCGTTCTGGAGCAGCCCGCAAGGCTGTTCT  7620
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CCACATCACTGCCCTTCGGGGCACTGATGTTCGTTCTGGAGCAGCCCGCAAGGCTGTTCT  7620
Qy  GGACTTGCAGAAGTGTGTGAGGGCAGTGAGATACCGAGTCATTATTGGCAAACTGTGAT  7680
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  GGACTTGCAGAAGTGTGTGAGGGCAGTGAGATACCGAGTCATTATTGGCAAACTGTGAT  7680
Qy  AGTTCCAAGGAGGAGGTCTTCGTGAAGACCAGGGAAGAAAACAACAAAGAAAACAACCCCCAAG  7740
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  AGTTCCAAGGAGGAGGTCTTCGTGAAGACCAGGGAAGAAAACAACAAAGAAAACAACCCCCAAG  7740
Qy  GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAGATGTACGGTCAGGT  7800
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAGATGTACGGTCAGGT  7800
Qy  TGCTCCTGACGTAGTTGAAAGCTGTCATGGGAGATGGCATTCACACCCATTCCACGTAC  7860
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TGCTCCTGACGTAGTTGAAAGCTGTCATGGGAGATGGCATTCACACCCATTCCACGTAC  7860
Qy  CCGTGTCAAGGCTCTGTTGTCGATGTCGGTCGCGGAGCCCGATGCAGTCGGAGCCCACATGCCGGATAC  7920
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CCGTGTCAAGGCTCTGTTGTCGATGTCGGTCGCGGAGCCCGATGCAGTCGGAGCCCACATGCCGGATAC  7920
Qy  AGTGTGTTTGACAGTACCATCACACACCCGGAGATATCATGTGGAGACAGACATCTACTC  7980
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  AGTGTGTTTGACAGTACCATCACACACCCGGAGATATCATGTGGAGACAGACATCTACTC  7980
Qy  AGCAGCTAAACTCAGTGACCAACACCCGAGCTGGCATTCACCAACCATTGCGAGGCAGTTATA  8040
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  AGCAGCTAAACTCAGTGACCAACACCCGAGCTGGCATTCACCAACCATTGCGAGGCAGTTATA  8040
Qy  CGCTGAGGACCGACGATGATCGCTTATGCGAGAGATCCGATATCGTAGGTGTAGGTC  8100
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CGCTGAGGACCGACGATGATCGCTTATGCGAGAGATCCGATATCGTAGGTGTAGGTC  8100
```

```
Db  TTCCGGGGTCTCATACTACCTTCAACAGTTGACCTGCTGCTGGCTGAAGGTAAATGC  8160
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TTCCGGGGTCTCATACTACCTTCAACAGTTGACCTGCTGCTGGCTGAAGGTAAATGC  8160
Qy  TGCAGCCGAACAGGCTGCATGAAGAACCCTGCTTCCTTATTGCGGCGATGATTGCAC  8220
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TGCAGCCGAACAGGCTGCATGAAGAACCCTGCTTCCTTATTGCGGCGATGATTGCAC  8220
Qy  CGTAATTGGAAGAGCGCCGGAGCAGATGCGACAAACAAGCAATGGCTGTGTCTTTGCTAG  8280
    ||||||||||||||||||||||||||||||||||||  |||||||||||
Db  CGTAATTGGAAGAGCGCCGGAGCAGATGCGACAGACAAACAAGCAATGGCTGTGTCTTTGCTAG  8280
Qy  CTGGATGAAGGTGATGGGTGCCACCACAGATTGTGTGCCTCAACCCAAATACAGTTTGGA  8340
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CTGGATGAAGGTGATGGGTGCCACCACAGATTGTGTGCCTCAACCCAAATACAGTTTGGA  8340
Qy  AGAATAACATGCTCATCAAATGTACCTCGGAATTACCAAAAGTGGCAAGCCTTA  8400
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  AGAATAACATGCTCATCAAATGTACCTCGGAATTACCAAAAGTGGCAAGCCTTA  8400
Qy  CTACTTTCTTACAAGACATCCTCGTATCCCCCTTGGCAGGTGCTCTCGCCGAGGGTCTGGG  8460
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CTACTTTCTTACAAGACATCCTCGTATCCCCCTTGGCAGGTGCTCTCGCCGAGGGTCTGGG  8460
Qy  ATACAACCCAGTGCTGCGTGGATTGGGTATCTAATACATCACTACCATGTTGTGGGT  8520
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  ATACAACCCAGTGCTGCGTGGATTGGGTATCTAATACATCACTACCATGTTGTGGGT  8520
Qy  TAGCCGTGTGTTGGCTGTCCATTTCATGGAGACAGATGCTCTTTGAGGACAAACTTCCCGA  8580
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TAGCCGTGTGTTGGCTGTCCATTTCATGGAGACAGATGCTCTTTGAGGACAAACTTCCCGA  8580
Qy  GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG  8640
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG  8640
Qy  CATCATTGCTGTGTGCACGGTATTGAGGCTTCTCGGTGGTGTGCGCTACCACCAACGCTGA  8700
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CATCATTGCTGTGTGCACGGTATTGAGGCTTCTCGGTGGTGTGCGCTACCACCAACGCTGA  8700
Qy  GATCCTCAGAGTTTCCAATCACTAACAGACATGACCATGGCTCCCGAGCCTGGGCG  8760
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  GATCCTCAGAGTTTCCAATCACTAACAGACATGACCATGGCTCCCGAGCCTGGGCG  8760
Qy  AAAGAAAGCCAGGGCGTCCTCGCGTCCTCGCGGAGCACCGTGGCGGAGCACGCGCAAAATT  8820
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  AAAGAAAGCCAGGGCGTCCTCGCGTCCTCGCGGAGCACCGTGGCGGAGCACGCGCAAAATT  8820
Qy  GGCTCGCTTCCTTCTCTCGGCCATGCAGGTCTCTAGACCTTCGGAGATTTGGATAAGACGAG  8880
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  GGCTCGCTTCCTTCTCTCGGCCATGCAGGTCTCTAGACCTTCGGAGATTTGGATAAGACGAG  8880
Qy  CGTGGCTCGGCAACCACTTTCAATTATTGTGATGTTTACTCCCGGAGGGGATGTGTT  8940
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CGTGGCTCGGCAACCACTTTCAATTATTGTGATGTTTACTCCCGGAGGGGATGTGTT  8940
Qy  TATTACACCCACAGAGAAGATTGCAGAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTTT  9000
    |||| ||||| |||||||||||||||||||||||||  |||||||| ||||||||
Db  TGTTACACCACAGAGAAGATTGCAGAGTTTCTTGTGAAGTATTTGGCTGTCATTGTTTT  9000
Qy  TGCCCTAGGGCTCATTGCTGTTGGATTGGATTAGCCATCAGTGAACCCCCAAATTCAAAATTAA  9060
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TGCCCTAGGGCTCATTGCTGTTGGACTAGCCATCAGTGAACCCCCAAATTCAAAATTAA  9060
Qy  CTAACAG----TTTTTTTTTTTTTTTTTTTTTTTAGGGCAACAGGGGAGACCCCC  9116
    |||||        |||||||||||||||||||||||||||
Db  TTTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCCGCAACAGGGGAGACCCCC  9120
Qy  GGGCTTAACGACCCCGC  9133
    |||||||||||||||||
Db  GGGCTTAACGACCCCGC  9137
```

RESULT 5
US-08-488-446-393
; Sequence 393, Application US/08488446
; Patent No. 6558898
GENERAL INFORMATION:
  APPLICANT: JOHN N. SIMONS
,,   : TAMI J. PILOT-MATIAS
,,   : GEORGE J. DAWSON
,,   : GEORGE G. SCHLAUDER
,,   : SURESH M. DESAI
,,   : THOMAS P. LEARY
,,   : ANTHONY SCOTT MUERHOFF
,,   : JAMES C. ERKER
,,   : SHERI L. BUIJK
,,   : ISA K. MUSHAHWAR
  TITLE OF INVENTION: NON-A, NON-B. NON-C, NON-D, NON-E HEPATITIS
  TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
  NUMBER OF SEQUENCES: 716
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
    STREET: 100 ABBOTT PARK ROAD
    CITY: ABBOTT PARK
    STATE: IL
    COUNTRY: USA
    ZIP: 60064-3500
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/488,446
    FILING DATE:
    CLASSIFICATION:
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US/08/424,550
    FILING DATE:
  ATTORNEY/AGENT INFORMATION:
    NAME: POREMBSKI, PRISCILLA E.
    REGISTRATION NUMBER: 33,207
    REFERENCE/DOCKET NUMBER: 5527.PC.01
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 708-937-6365
    TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 393:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 9143 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: linear
  MOLECULE TYPE: DNA (genomic)
  FEATURE:
    NAME/KEY: 5'UTR
    LOCATION: 1..445
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 446..9037
  FEATURE:
    NAME/KEY: 3'UTR
    LOCATION: 9038..9143
US-08-488-446-393

Query Match   96.4%;   Score 9059.8;   DB 4;   Length 9143;
Best Local Similarity   99.6%;   Pred. No. 0;
Matches 9096;   Conservative   0;   Mismatches   37;   Indels   4;   Gaps   1;

Qy   1 ACCACAAACACTCCAGTTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCACAAACACTCCAGTTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG 60

Qy   61 CAGGGCGTGGGGGATTCCCCTGCCCGTCTGCGAAGAGGTGGAGCCAACCACCTTAGTAT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CAGGGCGTGGGGGATTCCCCTGCCCGTCTGCGAAGAGGTGGAGCCAACCACCTTAGTAT 120

Qy   121 GTAGGCGGCGGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTAGGCGGCGGGGACTCATGACGCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC 180

Qy   181 CCTGATGGCGTTCATGGGTTCGGTGGTGGTGGCGCTTAGGCAGCCTCCACGCCCACCA 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCTGATGGCGTTCATGGGTTCGGTGGTGGTGGCGCTTAGGCAGCCTCCACGCCCACCA 240

Qy   241 CCTCCCAGATAGACGCGGGCGCACTGTAGGGAAGACCGGGGAGACCGGTCACTACCAAGGACG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCCCAGATAGACGCGGGCGCACTGTAGGGAAGACCGGGGAGACCGGTCACTACCAAGGACG 300

Qy   301 CAGACCTCCTTTTGAGTATCACGGCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTGT 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGACCTCCTTTTGAGTATCACGGCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTGT 360

Qy   361 TGGGATGGTTAGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TGGGATGGTTAGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT 420

Qy   421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTCTACTCAAACAAGTCCTGTACC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTCTACTCAAACAAGTCCTGTACC 480

Qy   481 TGGCCCCAGAACGCCAAGAACAAGCAGACGCAGGCTTCATATCCTGTGTCCATTAAAAC 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TGGCCCCAGAACGCCAAGAACAAGCAGACGCAGGCTTCATATCCTGTGTCCATTAAAAC 540

Qy   541 ATCTGTTGAAAGGGACAACGAGCAAAGCCAAAGTCCAGCGCGATGCTCGGCCTCGTAA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 ATCTGTTGAAAGGGACAACGAGCAAAGCCAAAGTCCAGCGCGATGCTCGGCCTCGTAA 600

Qy   601 TTACAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTACAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC 660

Qy   661 TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTCGCAATCTTGGAATCCTTCTGGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTCGCAATCTTGGAATCCTTCTGGA 720

Qy   721 TTACCCTTTGGGGTGGATTGGATGTTACAACTCACACACTTCAGTAGGCCGCTGGT 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TTACCCTTTGGGGTGGATTGGATGTTACAACTCACACACTTCAGTAGGCCGCTGGT 780

Qy   781 GGCAGGAGCGGTCGTTCGACCAGTCTGCCAGATAGTACGCTTGCTGGAGGATGGAGTCAA 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GGCAGGAGCGGTCGTTCGACCAGTCTGCCAGATAGTACGCTTGCTGGAGGATGGAGTCAA 840

Qy   841 CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTTGTGGTATGTCTCTATCTTTGGCCTG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTTGTGGTATGTCTCTATCTTTGGCCTG 900

Qy   901 TCCCTGTAGTGGGCGGCGGGGGTCACTGACCGACCCAGACACAAATACCACAATTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCCTGTAGTGGGCGGCGGGGGTCACTGACCGACCCAGACACAAATACCACAATTG 960

Qy   961 CTGCCAGCGTAATCAGGTTATCTATTGTCTCCTTCCACTTGCCTACACGAGCCTGGTTG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CTGCCAGCGTAATCAGGTTATCTATTGTCTCCTTCCACTTGCCTACACGAGCCTGGTTG 1020

Qy   1021 TGTGATCTGCGGACGAGCTGCTGGGTTCCCGCCAATCCGTACATCTCACCCCTTCCAA 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TGTGATCTGCGGACGAGCTGCTGGGTTCCCGCCAATCCGTACATCTCACCCCTTCCAA 1080

Qy   1081 TTGGACTGGCACGGACTCCTTCTTGGCTGACCACATTGATTTTGTTATGGGCGCTCTTGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TTGGACTGGCACGGACTCCTTCTTGGCTGACCACATTGATTTTGTTATGGGCGCTCTTGT 1140

Qy   1141 GACCTGTGACGCCCTGACATTGGTGAGTTGTGTGCGTGTATTAGTCGGTGACTG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTGTGACGCCCTGACATTGGTGAGTTGTGTGCGTGTATTAGTCGGTGACTG 1200

This page consists of a DNA sequence alignment listing. Each block aligns a query sequence (Qy) against a database sequence (Db) with match bars (|) between them.

Left column (positions 1201–2280):

```
Qy  1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGA  1260

Qy  1261  AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGTGATGGCCGGCCAAGGT  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGTGATGGCCGGCCAAGGT  1320

Qy  1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT  1380

Qy  1381  GTTTAGCAGTGTACACTACCTGGCGGTTGGCGCGTTGGCGCTCTCGGGGCAA  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  GTTTAGCAGTGTACACTACCTGGCGGTTGGCGCGTTGGCGCTCTCGGGGCAA  1440

Qy  1441  GTGGTATCAGTTGCTCCTAGGCGTTATGCTTTACATAGAAGCGACCTCTGAAACCCCAT  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  GTGGTATCAGTTGCTCCTAGGCGTTATGCTTTACATAGAAGCGACCTCTGAAACCCTAT  1500

Qy  1501  CAGGGGTGCCCACTGGATGCTCAATAGCTAGTTTGCTCGCCTTTGATGATACCATGTCC  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  CAGGGGTGCCCACTGGATGCTCAATAGCTAGTTTGCTCGCCTTTGATGATACCATGTCC  1560

Qy  1561  TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC  1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC  1620

Qy  1621  CAGGGCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTG  1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  CAGGGCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTG  1680

Qy  1681  TGCGAGGGGATGTGTATGGCTAAATTCAAAATAACAACATGGGGTTGCTGCCGTATTCGCAA  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  TGCGAGGGGATGTGTATGGCTAAATTCAAAATAACAACATGGGGTTGCTGCCGTATTCGCAA  1740

Qy  1741  TGTGCCATGTCACTACTGGGCACTATGGGCGTGTGGAACGACACTCGCACAACACTTA  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  TGTGCCATGTCACTACTGGGCACTATGGGCGTGTGGAACGACACTCGCACAACACTTA  1800

Qy  1801  CGAAGCAGCATGCCGTGTAACAACCGATGGCAACACGGCACAACGGCTCAGCCGCCTGAA  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  CGAAGCAGCATGCCGTGTAACAACCGATGGCAACACGGCACAACGGCTCAGCCGCCTGAA  1860

Qy  1861  ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC  1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC  1920

Qy  1921  AGGCCATTGTATTTGAGGGATCAGATATAGTTTACTTTTATGACCCTGTGAA  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  AGGCCATTGTATTTGAGGGATCAGATATAGTTTACTTTTATGACCCTGTGAA  1980

Qy  1981  TTCCACTCTCTACCACCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCTGTACG  2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  TTCCACTCTCTACCACCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCTGTACG  2040

Qy  2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTCACACAGG  2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTCACACAGG  2100

Qy  2101  ATTGATCACCAAAGACAAAGCCTGGAAAAAAATTATCAGGTCTTATATTCCGCCACGGTGC  2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  ATTGATCACCAAAGACAAAGCCTGGAAAAAAATTATCAGGTCTTATATTCCGCCACGGTGC  2160

Qy  2161  TTTGTCTCTTACGGAGGTTACCACCAGAGGCCGTGGCTAATTCTGTTGGGGTTGTGTGG  2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TTTGTCTCTTACGGAGGTTACCACCAGAGGCCGTGGCTAATTCTGTTGGGGTTGTGTGG  2220

Qy  2221  CAGCAAGTATCTTATTTTAGCCTAAGTCTCTGTTACTTGTCCCTTGTTTGTTGGGGCGCTTC  2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  CAGCAAGTATCTTATTTTAGCCTAAGTCTCTGTTACTTGTCCCTTGTTTGTTGGGGCGCTTC  2280
```

Right column (positions 2281–3420):

```
Qy  2281  TGGTTACCCTTTGCGTCCTGTTGCTCCATCCCAGTCGTGTATCTCCAAGCTGGCTGGGGATGT  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TGGTTACCCTTTGCGTCCTGTTGCTCCATCCCAGTCGTGTATCTCCAAGCTGGCTGGGGATGT  2340

Qy  2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTTGCTTGATTTCTTCATCTGTTGATTTCTCATCTGTATCTCCG  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTTGCTTGATTTCTTCATCTGTTGATTTCTCATCTGTATCTCCG  2400

Qy  2401  CTGCAGGCTACGTTATGCTGCCCCATTGTTGTGCCCATGGCTCGCGGGCTTGCCCCT  2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  CTGCAGGCTACGTTATGCTGCCCCATTGTTGTGCCCATGGCTCGCGGGCTTGCCCCT  2460

Qy  2461  AACTTTCTTGTTGCAGCAGCTGCTGCCCAACCAGATTATGACTGGTGGTGCCGACTGCT  2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  AACTTTCTTGTTGCAGCAGCTGCTGCCCAACCAGATTATGACTGGTGGTGCCGACTGCT  2520

Qy  2521  AGTGGGCAGGGTAGTTTGTGTGGGCCCGCCGTAACCGTGGTCACCGCATAGCCTGCTTGT  2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  AGTGGGCAGGGTAGTTTGTGTGGGCCCGCCGTAACCGTGGTCACCGCATAGCCTGCTTGT  2580

Qy  2581  AGGTCCTTGGCCTCTGCTAGGCGTTTAACCCTCTTGTCATTTGGTTACGCCCTATCT  2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  AGGTCCTTGGCCTCTGCTAGGCGTTTAACCCTCTTGGCATTTGGTTACGCCCTATCT  2640

Qy  2641  TTTGATACCGAGATAATTGGAGGGCTGACAATACCACCCTGTAGTAGCCATTAGTGTCAT  2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  TTTGACACCGAGATAATTGGAGGGCTGACAATACCACCCTGTAGTAGCCATTAGTGTCAT  2700

Qy  2701  GTCTCGTTTTGGCTTCTTGCTCACTGTTACCTCGGTTGTTAGTTAACTCCTATCT  2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  GTCTCGTTTTGGCTTCTTGCTCACTGTTACCTCGGTTGTTAGTTAACTCCTATCT  2760

Qy  2761  TTGGCAACGTTGGGAGAATTGGTTTTGGAACGGTTGCTGCCTATTCGCAA  2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  TTGGCAACGTTGGGAGAATTGGTTTTGGAACGGTTGCTGCCTATTCGCAA  2820

Qy  2821  TGTGCTCGGTTTGTTGTTCCCCGGTGCGCACATATGACGCGGCTGGTGACTTTCTCGTGTGTCA  2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  TGTGCTCGGTTTGTTGTTCCCCGGTGCGCACATATGACGCGGCTGGTGACTTTCTCGTGTGTCA  2880

Qy  2881  CGTAGCTCTTCTTAAACATCCAGTGCGCAGCATCGTTCTTTGGGACTGACTCTAGGGT  2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  CGTAGCTCTTCTTAAACATCCAGTGCGCAGCATCGTTCTTTGGGACTGACTCTAGGGT  2940

Qy  2941  TAGGGCCCATAGAAGTGGTGCGTCTCGGAAAGTGTCATGCTTGGTATTCTCATTATGT  3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941  TAGGGCCCATAGAAGTGGTGCGTCTCGGAAAGTGTCATGCTTGGTATTCTCATTATGT  3000

Qy  3001  TCTTAAGTTTTCCTCTAGTGTTGGTGAGAATGGTGTTTTTCATATAAGCACTTGCA  3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  TCTTAAGTTTTCCTCTAGTGTTGGTGAGAATGGTGTTTTTCATATAAGCACTTGCA  3060

Qy  3061  TGGTGATGGTCTTGCCTGCTAATGATTTGCCTCGAAACTACCGCCCATTTTTTCCC  3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061  TGGTGATGGTCTTGCCTGCTAATGATTTGCCTCGAAACTACCGCCCATTTTTTCCC  3120

Qy  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGGCTTTACAGTGATGTGGGGGCACAC  3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGGCTTTACAGTGATGTGGGGGCACAC  3180

Qy  3181  GGTTGATGGTTGCCCGGTTCGCCGGTGTTTTCGGAGGGTTAGCTAT  3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181  GGTTGATGGTTGCCCGGTTCGCCGGTGTTTTCGGAGGGTTAGCTAT  3240

Qy  3241  GCCGCCAGATGGGTGGGCCATTACCGGCAGTGTCTCTCGACCGGTGG  3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241  GCCGCCAGATGGGTGGGCCATTACCGGCAGTGTCTCTCGACCGGTGG  3300

Qy  3301  CACGCTGTCAGCGATGCAGTGGTCATGACTTGACTGGAAC  3360
          ||||||||||||||||||||||||||||||||||||||||
Db  3301  CACGCTGTCAGCGATGCAGTGGTCATGACTTGACTGGAAC  3360

Qy  3361  TATCTTCAGATTAGGATCTCTGGCCACTAGCTACACATGGGATTTGTTTGTGCAACGTGTT  3420
```

Right column:

Db 4441 ATGTGACATCTCAAAAATCCCTGAGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500

Qy 4501 TACAGGGTACACTGGTGACTTTGATTCGTGCTGATGACTGGCAGCGTCATGTAGAAGGCAC 4560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4501 TACAGGGTACACTGGTGACTTTGATTCGTGCTGATGACTGGCAGCGTCATGTAGAAGGCAC 4560

Qy 4561 ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTGTTCGTGTGTGCGGGGTTTCAGC 4620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4561 ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTCGTTCGTGTGTGCGGGGTCTCAGC 4620

Qy 4621 AATAGTTAAAGGCCAGCGTAGGGCGTGGGAGAGTCGGCATATACTACTTA 4680
          |||||||||||||||||||||||||||||||||||||||||||||
Db 4621 AATAGTTAAAGGCCAGCGTAGGGCGTGGGAGACTCGGCATATACTACTTA 4680

Qy 4681 TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 4681 TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740

Qy 4741 CGACGCAGCCAAGGCATGTATGGTTTGTCATCAACAGAGCTCAAACTATTCTGGACAC 4800
          |||||||||||||||||||||||||||||||||||||||||||||||
Db 4741 CGACGCAGCCAAGGCATGTATGGTTTGTCATCAACAGAGCTCAAACTATTCTGGACAC 4800

Qy 4801 CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT 4860
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 4801 CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT 4860

Qy 4861 CTTTTCTATGGTCAACCCGAACCTTCATTGTCAATACTGCAAAAAGAACTGCTGACAA 4920
          ||||||||||||||||||||||||||||||||||||||||||||||
Db 4861 CTTTTCTATGGTCAACCCGAACCTTCATTGTCAATACTGCAAAAAGAACTGCTGACAA 4920

Qy 4921 TTATGTTTTGTTGACTGCAGCCAACTACAACTGTCATCAGTAGTGCTATGCTCCC 4980
          ||||||||||||||||||||||||||||||||||||||||||||
Db 4921 TTATGTTTTGTTGACTGCAGCCAACTACAACTGTCATCAGTAGTGCTATGCTCCC 4980

Qy 4981 CAATGACGCACCACGGTGCAGGGTGCGCAGGGAGCCCGGCTTGGGAAAAACCTTGTCTGTG 5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4981 CAATGACGCACCACGGTGCAGGGTGCGCAGGGAGCCCGGCTTGGGAAAAACCTTGTCTGTG 5040

Qy 5041 GCGCTTGGACGGCGGCTGACGCGCTGCGTACTTCCTGGGACAGCCCGACTCCCGTTGGCGT 5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5041 GCGCTTGGACGGCGGCTGACGCGCTGCGTACTTCCTGGGACAGCCCGACTCCCGTTGGCGT 5100

Qy 5101 AATGTGCTTCACTGAAGTCAATACTTCCTTGGAGGCCATGTGTGCTGCAATTGACAAGCTGCAT 5160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5101 AATGTGCTTCACTGAAGTCAATACTTCCTTGGAGGCCATGTGTGCTGCGATAAGCTGAA 5160

Qy 5161 GGCTATGGCTTATCTAGCCATTGCAGCACTTTTGGCGCCCACTTGTGTGCGGCGTTGCTGGTC 5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5161 GGCTATGGCTTATCTAGCCATTGCAGCACTTTTGGCGCCCACTTGTGTGCGGCGTTGCTGGTC 5220

Qy 5221 TATTACATCAGTCCCTACCGGTGCTACTGGTGTTGACGGAAGAAGAAATCGT 5280
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db 5221 TATTACATCAGTCCCTACCGGTGCTACTGGTGTTGACGGAAGAAGAAATCGT 5280

Qy 5281 GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGTGTGCTGCAATTGACAAGCTGAA 5340
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db 5281 GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGTGTGCTGCGATAAGCTGAA 5340

Qy 5341 GAGTACAATCACCACAACTAGTCTTTCACATTGGAAACGCCCTTGAAAAACTTAACAC 5400
          |||||||||||||||||||||||||||||||||||||||||||||
Db 5341 GAGTACAATAACCACAACTAGTCCTTTCACATTGGAAACGCCCTTGAAAAACTTAACAC 5400

Qy 5401 CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGTTTAGT 5460
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5401 CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGGCTTAGT 5460

Qy 5461 CACTTTACCTGACAATCCCTTTGCTTTGCTTTGCTTTGCCGGGTATTACTAC 5520
          ||||||||||||||||||||||||||||||||||||||||||
Db 5461 CACTTTACCTGACAATCCCTTTGCATCGTGTTTGCTTTGCCGGGTATTACTAC 5520

Qy 5521 CCCACTACCTCACAAGATCAAAAATGTTCCTGTCATTATTTGGAGGCGCAATTGCCGTCCAA 5580
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5521 CCCACTACCTCACAGATCAAAAATGTTCCTGTCATTATTTGGAGGCGCAATTGCCGTCCAA 5580

Left column:

Db 3361 ... |||||||||||||||||||||||||||||||||||| 3420
           TATCTTCAGATTAGGATCTGGCCACTAGTCACATGGGATTTGTTTGTGACAACGTGTT

Qy 3421 GTATACTGCTCACCATGGCAGCAAGGGGCGCCGGTTGGCTCATCCCACAGGCTCTATACA 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3421 GTATACTGCTCACCATGGCAGCAAGGGGCGCCGGTTGGCTCATCCCACAGGCTCCATACA 3480

Qy 3481 CCCAATAACCGTTGACGCGGCTAATGACCATCTATCAACCACCATGTGGAGCTGG 3540
          |||||||||||||||||||||||||||||||||||||||||||||
Db 3481 CCCAATAACCGTTGACGCGGCTAATGACCATCTATCAACCACCATGTGGAGCTGG 3540

Qy 3541 GTCCCTTACTCGGTCGTGCTCTTCGGGGAGACCAAGGGGTATCGGTAACACGACTGGGGTC 3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3541 GTCCCTTACTCGGTCGTGCTCTTCGGGGAGACCAAGGGGTATCGGTAACACGACTGGGGTC 3600

Qy 3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGCGGGGCCCTTCCAT 3660
          |||||||||||||||||||||||||||||||||||||||||||||||
Db 3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGCGGGGCCCCTTCCAT 3660

Qy 3661 GGCTGTGTGCCAAGGGTTCTTCAGTGCCCGATTCTGTGCTCCTCCGGGCATGTTATTGG 3720
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db 3661 GGCTGTGTGCCAAGGGTTCTTCAGTGCCCGATTCTGTGCTCCTCCGGGCATGTTATTGG 3720

Qy 3721 GATGTTCACCCGCTGCTAGAAATTCTGCGTCAGTCAGTTCAGGGTTAGGCCGTT 3780
          ||||||||||||||||||||||||||||||||||||||||||
Db 3721 GATGTTCACCCGCTGCTAGAAATTCTGCGTCAGTCAGTTCAGGGTTAGGGCCGTT 3780

Qy 3781 GGTGTGTGCTGGATACCAATCCCCAGTACACAGACATGCCACTCTGATACAAAACCTAC 3840
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db 3781 GGTGTGTGCTGGATACCAATCCCCAGTACACAGACATGCCACTCTGATACAAAACCTAC 3840

Qy 3841 TGTGCCTAACAGTATTCAGTGCAAATTTAATTGCCCCACTGAGCGGCAAGTCAAC 3900
          ||||||||||||||||||||||||||||||||||||||||||
Db 3841 TGTGCCTAACAGTATTCAGTGCAAATTTAATTGCCCCACTGGCGAGCGGCAAGTCAAC 3900

Qy 3901 CAAATTACCACTTCTTCTACAATGCAGGAAGAATGATGAGGTCTTGGTCCTAAATCCCAGTGT 3960
          |||||||||||||||||||||||||||||||||||||||||||
Db 3901 CAAATTACCACTTCTTCTACAATGCAGGAAGAATGATGAGGTCTTGGTCCTAAATCCCAGTGT 3960

Qy 3961 GGCTACAACAGCATCAATGCCAAAGTACAGTACGCGACGTACGCGCGTGAATCCAAATTG 4020
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db 3961 GGCTACAACAGCATCAATGCCAAAGTACAGTACGCGACGTACGCGCGTGAATCCAAATTG 4020

Qy 4021 CTATTTTAATGCAAATGTACCAACACAGGGGCTTCACTTACTTCACTTACGTACACCAC 4080
          ||||||||||||||||||||||||||||||||||||||||||||||
Db 4021 CTATTTTAATGCAAATGTACCAACACAGGGGCTTCACTTACTTCACTTACGTACACCAC 4080

Qy 4081 GTACCTGACCGGAGCATGTTCCCGGAACTATGATGTAATCATTGTGACGGAATGCCATGC 4140
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db 4081 GTACCTGACCGGAGCATGTTCCCGGAACTATGATGTAATCATTGTGACGGAATGCCATGC 4140

Qy 4141 TACCGATGCAACCACCGTGTTGGGCATTGGGAAGGTCCTTACCGAAGCTGCATCCCAAAAA 4200
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db 4141 TACCGATGCAACCACCGTGTTGGGCATTGGGAAGGTTCTAACCGAAGCTGCATCCCAAAAA 4200

Qy 4201 TGTTAGGCTAGTTGCTCAACGAGTTAGCTCGAAAGGAATAACAGCTGTCTCTTACTTACTAGGGG 4260
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4201 TGTTAGGCTAGTTGCTCAACGAGTTAGCTCGAAAGGAATAACAGCTGTCTCTTACTTACTAGGGG 4260

Qy 4261 CAACATAACTGAGATTCAATTAACCGGCACTTCCTGGAAAGGCACTTCCTGGAAAAAAGAT 4320
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 4261 CAACATAACTGAGATTCAATTAACCGGCACTTCCTGGAAAGGCACTTCCTGGAAAAAAGAT 4320

Qy 4321 TAAGGAGAAAATCTGAAGAAAGGGAGACACTGTATCTTTGAGGCTACCAAAAAACACTG 4380
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db 4321 TAAGGAGAAAATCTGAAGAAAGGGAGACACCTTATCTTTGAGGCTACCAAAAAACACTG 4380

Qy 4381 TGATGAGCTTGCTCAACGAGTTAGCTCGAAAGGAATAACAGCTGTCTCTTACTATAGGGG 4440
          ||||||||||||||||||||||||||||||||||||||||||||||
Db 4381 TGATGAGCTTGCTCAACGAGTTAGCTCGAAAGGAATAACAGCTGTCTCTTACTATAGGGG 4440

Qy 4441 ATGTGACATCTCAAAAATCCCTGAGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500
          ||||||||||||||||||||||||||

```
Qy  5581  GCTTACGACGCTAGAGGCGACTGGCGTTCATGATGGCCGGGGCGTGCGGGAACAGCTCT  5640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581  GCTTACGACGCTAGAGGCGACTGGCGTTCATGATGGCCGGGGCGTGCGGGAACAGCTCT  5640

Qy  5641  TGGTACATGGACATCGGTGGGTTTGCTTTGACATGTAGGCGGCTATGCTGCCGCCTC  5700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641  TGGTACATGGACATCGGTGGGTTTGCTTTGACATGTAGGCGGCTATGCTGCCGCCTC  5700

Qy  5701  ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGTGAGTGGCCCACTATGGATCAGCT  5760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701  ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGTGAGTGGCCCACTATGGATCAGCT  5760

Qy  5761  TGCTGGTTTAGTTGACTTACTCCGGCGTTCAATCCGGCCGGAGGAGTGTGGGCGTCTTGTCAGC  5820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761  TGCTGGTTTAGTTGACTTACTCCGGCGTTCAATCCGGCCGGAGGAGTGTGGGCGTCTTGTCAGC  5820

Qy  5821  TTGTGCAATGTTGCTTGACAACAGCAGGGCCAGATCACTGGCCCAACAGACTTCTTAC  5880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821  TTGTGCAATGTTGCTTGACAACAGCAGGGCCAGATCACTGGCCCAACAGACTTCTTAC  5880

Qy  5881  TATGCTTGCTAGGAGCAACACTGTATGTGAATGAGTACTTTATTGCCACTCGTGACATCCG  5940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881  TATGCTTGCTAGGAGCAACACTGTATGTGAATGAGTACTTTATTGCCACTCGTGACATCCG  5940

Qy  5941  CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCCTGGAGTGTCATATCAGCTTGCAT  6000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941  CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCCTGGAGTGTCATATCAGCTTGCAT  6000

Qy  6001  CCGTTGGCTCCACACCCCGACGGAGATGATTGCGGCTCATTGGCTGGGGTCTAGAGAT  6060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001  CCGTTGGCTCCACACCCCGACGGAGATGATTGCGGCTCATTGGCTGGGGTCTAGAGAT  6060

Qy  6061  TTGGCAGTATGTGTGCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGCTGGAGTTCA  6120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061  TTGGCAGTATGTGTGCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGCTGGAGTTCA  6120

Qy  6121  GAGCATGGTTAACATTCCTGGTTGTGTCCTTTCTACAGCTGCCAGAAGAGGGGTACAAGGGCCC  6180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121  GAGCATGGTTAACATTCCTGGTTGTGTCCTTTCTACAGCTGCCAGAAGAGGGGTACAAGGGCCC  6180

Qy  6181  CTGGAATGGATCAGGTAGGCGTATGCTCCAAGCACGCGTCTGAACTCATCTTTTC  6240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181  CTGGAATGGATCAGGTAGGCGTATGCTCCAAGCACGCGTCTGAACTCATCTTTTC  6240

Qy  6241  TGTTGAGAAATGGTTTTGCAAAACTTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241  TGTTGAGAAATGGTTTTGCAAAACTTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300

Qy  6301  AGGGGCTGTTCCAGTCAACGCTAGGCGTGTGTGGGTCGGCTAGACGAAATATGGGGGATTG  6360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301  AGGGGCTGTTCCAGTCAACGCTAGGCGTGTGTGGGTCGGCTAGACGAAATATGGGGGATTG  6360

Qy  6361  GACTAGTCTTGTCGTCAATTATTGGCGTCAATGTCTGTTTCACCCAGTGCCCCCAAC  6420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361  GACTAGTCTTGTCGTCAATTATTGGCGTCAATGTCTGTTTCACCCAGTGCCCCCAAC  6420

Qy  6421  TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTTGTTTCACCCAGGTAAAGTTAGGTGGGAGA  6480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421  TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTTGTTTCACCCAGGTAAAGTTAGGTGGGAGA  6480

Qy  6481  CTTGAGAGCTGCAGTGCCGTGACGCGTACAGGTTCAGTGTTATCGGGTTAGGTGAGCCCAA  6540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481  CTTGAGAGCTGCAGTGCCGTGACGCGTACAGGTTCAGTGTTATCGGGTTAGGTGAGCCCAA  6540

Qy  6541  AACTCCTTGACGACATCTGCTTGCTGTTACGCGTCTGTTACGCGTCCGGACGTAAGGGTAAAACTGTTAA  6600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541  AACTCCTTGACGACATCTGCTTGCTGTTACGCGTCTGTTACGCGTCCGGACGTAAGGGTAAAACTGTTAA  6600

Qy  6601  GCTTCCCTTCCCGCGTTGACGGTCACACCTTGGTGCGCATGCGAACTTAATTTGCGTGA  6660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601  GCTTCCCTTCCCGCGTTGACGGTCACACCTTGGTGCGCATGCGAACTTAATTTGCGTGA  6660


Qy  6661  TGCACTTGAGACAAATGACTGTAATTCCACAAACAACACTCCTAGTGATGAGCCGCGAGT  6720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661  TGCACTTGAGACAAATGACTGTAATTCCATAAACAACACTCCTAGTGATGAGCCGCGAGT  6720

Qy  6721  GTCCGCTCTTGTTTCAAACGAGTTGCGCGTACAAACCAATTGCTTGAGGCAATTTC  6780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721  GTCCGCTCTTGTTTCAAACGAGTTGCGCGTACAAACCAATTGCTTGAGGCAATTTC  6780

Qy  6781  AGCTGGGGTTGACACCACCAAACTGCCAGCCCCCTCCATCGAAGAGGTAGTGTAAGAAA  6840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781  AGCTGGCGTTGACACCACCAAACTGCCAGCCCCCTCCATCGAAGAGGTAGTGTAAGAAA  6840

Qy  6841  GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCCTCCGAGATCCGTCCC  6900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841  GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCCTCCGAGATCCGTCCC  6900

Qy  6901  AGGAGTGTCATGTCCTGAAAGCTCGAACGTGAGCCGTTAGAAGAGTTCCTCAAACCT  6960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901  AGGAGTGTCATGTCCTGAAAGCTCGAACGTGAGCCGTTAGAAGAGTTCCTCAAACCT  6960

Qy  6961  CCCTCCTTCACCACCTGTCTACAGTTGGCCATGCCGATGCCCCTGTTGGGAGCGGGTGA  7020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961  CCCTTCTTCACCACCTGTCTACAGTTGGCCATGCCGATGCCCCTGTTGGGAGCGGGTGA  7020

Qy  7021  GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAACAGGCCGGAGGGCCCTGATGA  7080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021  GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAACAGGCCGGAGGGCCCTGATGA  7080

Qy  7081  TTTACCCAGTTACCCTCCCAAAAAGAGGAGGTCTCGAATGGTCAGCGAAAGTTGGTCGAC  7140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081  TTTACCCAGTTACCCTCCCAAAAAGAGGAGGTCTCGAATGGTCAGCGAAAGTTGGTCAAC  7140

Qy  7141  GGCTACAACCGCTTCAGCTGGTTACTGGCCCCCCGTACCCTAAGAGTAACCGGGGAAAGGA  7200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141  GACTACAACCGCTTCAGCTGGTTACTGGCCCCCCGTACCCTAAGAGTAACCGGGGAAAGGA  7200

Qy  7201  TTCCACTCAGTCAGCCCCGCCCAAACGGCCTACAAAAAGAGTGGGAAGAGTGAGTT  7260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201  TTTCACTCAGTCAGCCCCGCCCAAACGGCCTACAAAAAGAGTGGGAAGAGTGAGTT  7260

Qy  7261  TTCGTGCAGATGAGCTACACCTGACGACTGATGAGTTCTTCAAAACTGCTTCTAAAGT  7320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261  TTCGTGCAGATGAGCTACACCTGACGACTGATGAGTTAGCTTCAAAACTGCTTCTAAAGT  7320

Qy  7321  TCTGTCTGCAACTCGGGCCATCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTATGT  7380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7321  TCTGTCTGCAACTCGGGCCATCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTATGT  7380

Qy  7381  GACTGAGCCGCGGATGCGGAGCTTGAGAAGTCACTATTAATAGACAACCTCT  7440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381  GACTGAGCCGCGGATGCGGAGCTTGAGAAGTCACTATTAATAGACAACCTCT  7440

Qy  7441  GTTCCCCCATCATACACCAGAAGCAAGTTGGCTAAGGAAGAAAAGCTTCAAAAGTTGT  7500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441  GTTCCCCCATCATACACCAGAAGCAAGTTGGCTAAGGAAGAAAAGCTTCAAAAGTTGT  7500

Qy  7501  CGGTGTCATGTGGACTATGATGAGCAGTCACACGCCCTCTAAGTCTCTAAGTC  7560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7501  CGGTGTCATGTGGACTATGATGAGCAGTCACACGCCCTCTAAGTCTCTAAGTC  7560

Qy  7561  CCACATCACTGCCTTCGGGGACTGATGTTCGTTCTGGAGCAGCCCGCAAGGCTGTTCT  7620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7561  CCACATCACTGCCTTCGGGGACTGATGTTCGTTCTGGAGCAGCCCGCAAGGCTGTTCT  7620

Qy  7621  GGACTTGCAGAAGTGTGTCGAGGCAGTGAGATGAGTGCAAACTGTGAT  7680
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  7621  GGACTTGCAGAAGTGTGTCGAGGCAGTGAGATGAGTGCAAACTGTGAT  7680

Qy  7681  AGTTCCAAAGGAGGAGGTCTTGTGAAGAACCCCCAGAAACCAACAAAGAAAACCCCCAAG  7740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7681  AGTTCCAAAGGAGGAGGTCTTGTGAAGAACCCCCAGAAACCAACAAAGAAAACCCCCAAG  7740

Qy  7741  GCTTATCTCGTACCCCACCTTGAAATGAGTGTTGAGAAGATGTACTACGGTCAGGT  7800
```

RESULT 6
US-08-467-344A-390
; Sequence 390, Application US/08467344A
; Patent No. 6586568
GENERAL INFORMATION:
    APPLICANT: JOHN N. SIMONS
              TAMI J. PILOT-MATIAS
              GEORGE J. DAWSON
              GEORGE G. SCHLAUDER
              SURESH M. DESAI
              THOMAS P. LEARY
              ANTHONY SCOTT MUERHOFF
              JAMES C. ERKER
              SHERI L. BUIJK
              ISA K. MUSHAHWAR
    TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
                        REAGENTS AND METHODS FOR THEIR USE
    NUMBER OF SEQUENCES: 716
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: ABBOTT LABORATORIES   D377/AP6D
        STREET: 100 ABBOTT PARK ROAD
        CITY: ABBOTT PARK
        STATE: IL
        COUNTRY: USA
        ZIP: 60064-3500
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/467,344A
        FILING DATE: 07-Jun-1995
        CLASSIFICATION: <Unknown>
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: 08/424,550
        FILING DATE: <Unknown>
    ATTORNEY/AGENT INFORMATION:
        NAME: POREMBSKI, PRISCILLA E.
        REGISTRATION NUMBER: 33,207
        REFERENCE/DOCKET NUMBER: 5527.PC.01
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: 708-937-6365
        TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 390:
    SEQUENCE CHARACTERISTICS:
        LENGTH: 9143 base pairs
        TYPE: nucleic acid
        STRANDEDNESS: double
        TOPOLOGY: linear

```
;
;  MOLECULE TYPE: DNA (genomic)
;  SEQUENCE DESCRIPTION: SEQ ID NO: 390:
US-08-467-344A-390

       Query Match      96.4%;  Score 9059.8;  DB 4;  Length 9143;
       Best Local Similarity 99.6%;  Pred. No. 0;
       Matches 9096;  Conservative 0;  Mismatches 37;  Indels 4;  Gaps 1;
```

```
Qy    1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTTCGGAGCACCCCCCTAG   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTTCGGAGCACCCCCCTAG   60

Qy   61 CAGGGGCGTGGGGGGGGATTCCCCTGCCGTGCAGGAGGGTGGAGCCAAGGCCAACCACTTGACTTGATGGC  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CAGGGGCGTGGGGGGGGATTCCCCTGCCGTGCAGGAGGGTGGAGCCAAGGCCAACCACTTGACTTGATGGC  120

Qy  121 GTAGGCGGCGGCGGCGCCACTGTAGGAGCGGCGGCGCTCGGTGATGACAAGCGGCCAAGCTTGACTTGGTGGTGGC  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTAGGCGGCGGCGGCGCCACTGTAGGAGCGGCGGCGCTCGGTGATGACAAGCGGCCAAGCTTGACTTGGTGGTGGC  180

Qy  181 CCTGATGGGCGTTCATGGGTTCGGTGGTGGGCGGCGTTTAGGCAGCCTCCACGCCCACCA  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CCTGATGGGCGTTCATGGGTTCGGTGGTGGGCGGCGTTTAGGCAGCCTCCACGCCCACCA  240

Qy  241 CCTCCCAGATAGAGCGGCGGCCACTGTAGGAGGAAGACCCGGGAACCGGTCACTACCAAGGACG  300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 CCTCCCAGATAGAGCGGCGGCCACTGTAGGAGGAAGACCCGGGAACCGGTCACTACCAAGGACG  300

Qy  301 CAGACCTCTCTTTTGAGTATCACGCCTTCGGGAAGTAGTTGGGCAAGCCCACCTATATGTGT  360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGACCTCTCTTTTGAGTATCACGCCTTCGGGAAGTAGTTGGGCAAGCCCACCTATATGTGT  360

Qy  361 TGGGATGGTTGGGGTTAGCCATCCATACGTACTGCCTGATAGGGTCCTTGCCTGAGGGGAT  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 TGGGATGGTTGGGGTTAGCCATCCATACGTACTGCCTGATAGGGTCCTTGCCTGAGGGGAT  420

Qy  421 CTGGGAGTCTTCGTAGACCGTAGCACATGCCGTGTTATTTCTACTCAAACAAGTCCTGTACC  480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CTGGGAGTCTTCGTAGACCGTAGCACATGCCGTGTTATTTCTACTCAAACAAGTCCTGTACC  480

Qy  481 TGCGCCCAGAAACGCGCAAGAACGCGCAAGAACGCGCAAGCGCAAAGTCCAGGCCTTTGCCAGC  540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TGCGCCCAGAAACGCGCAAGAACGCGCAAGAACGCGCAAGCGCAAAGTCCAGGCCTTTGCCAGC  540

Qy  541 ATCTGTTGAAAGGGACAACGACAACGACGAGCAAAGTCCAGGCGGATGCTCGGCCTCGTAA  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 ATCTGTTGAAAGGGACAACGACAACGACGAGCAAAGTCCAGGCGGATGCTCGGCCTCGTAA  600

Qy  601 TTACAAAATTGCTGTATCCATGATGGCTTGCAGACATTGGCTTGGCCATAAGTCTCGCAATCCTTCTGA  660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 TTACAAAATTGCTGTATCCATGATGGCTTGCAGACATTGGCTTGGCCATAAGTCTCGCAATCCTTCTGA  660

Qy  661 TCATGGTTGGGACGGCCAAGACCCTCGCCATAAGTCTCGCAATCTTGGAATCCTTCTGA  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TCATGGTTGGGACGGCCAAGACCCTCGCCATAAGTCTCGCAATCTTGGAATCCTTCTGA  720

Qy  721 TTACCCTTTGGGGTGATTGGTGATGTTACAACTCACACACCTCTAGTAGGCCCGCTGGT  780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 TTACCCTTTGGGGTGATTGGTGATGTTACAACTCACACACCTCTAGTAGGCCCGCTGGT  780

Qy  781 GGCAGGAGCCGGTCGTTCGACCAGTCTGCCAGATAGTACGCCTGCTGGAGATGGAGTCAA  840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GGCAGGAGCCGGTCGTTCGACCAGTCTGCCAGATAGTACGCCTGCTGGAGATGGAGTCAA  840

Qy  841 CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTGTGTGTATCTGCTATCTTTGGCCTG  900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTGTGTGTATCTGCTATCTTTGGCCTG  900

Qy  901 TCCCTGTAGTGGGCGGGGGTCACTGACCCAGACACAAATACCACAATCCTGACCAATTG  960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TCCCTGTAGTGGGCGGGGGTCACTGACCCAGACACAAATACCACAATCCTGACCAATTG  960

Qy  961 CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCACTTGCCTACACGAGCCTGGTTG 1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCACTTGCCTACACGAGCCTGGTTG 1020
```

```
Qy  961 CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCACTTGCCTACACGAGCCTGGTTG 1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCACTTGCCTACACGAGCCTGGTTG 1080

Qy 1021 TGTGATCTGTCGCGGGACGAGTGCTGGGTCTCGGCCAATCCGTACATCTCACACCTTCCAA 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 TGTGATCTGTCGCGGGACGAGTGCTGGGTCTCGGCCAATCCGTACATCTCACACCTTCCAA 1080

Qy 1081 TTGGACTGGCACGGACTCCTTCTTGGCTGGCTGACCACATTGATTTTGTTATGGGCGTCTTGT 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 TTGGACTGGCACGGACTCCTTCTTGGCTGGCTGACCACATTGATTTTGTTATGGGCGTCTTGT 1140

Qy 1141 GACCTGTGACGCCCTTGACATTGTGAGTTGTGTGGCGTGTATTAGTCGGTGACTG 1200
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 GACCTGTGACGCCCTTGACATTGTGAGTTGTGTGGCGTGTATTAGTCGGTGACTG 1200

Qy 1201 GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTTGA 1260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTTGA 1260

Qy 1261 AGTGCCCACTGGAATAGATCCTCGGGTTCCTAGGGTTTATCGGGATGGCCGGCCAAGGT 1320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 AGTGCCCACTGGAATAGATCCTCGGGTTCCTAGGGTTTATCGGGATGGCCGGCCAAGGT 1320

Qy 1321 CGAGGCGTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGGTATTGGCGACTAT 1380
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 CGAGGCGTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGGTATTGGCGACTAT 1380

Qy 1381 GTTTAGCAGTGTACACTGCCGGTTGGCCGTCTCGATCTCATCTACTACTGCCTCTCGGGGCAA 1440
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 GTTTAGCAGTGTACACTGCCGGTTGGCCGTCTCGATCTCATCTACTACTGCCTCTCGGGGCAA 1440

Qy 1441 GTGGTATCAGTTGCTCTCCTAGCGTCTTACATAGAAGCGACCTCTGGAAACCCCAT 1500
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 GTGGTATCAGTTGCTCTCCTAGCGTCTTACATAGAAGCGACCTCTGGAAACCCTAT 1500

Qy 1501 CAGGGTGCCCACTGGATGCTAGCACATGCCGTTGCCCTTTGCTCGCCTTGATGATACCATGTCC 1560
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 CAGGGTGCCCACTGGATGCTAGCACATGCCGTTGCCCTTTGCTCGCCTTGATGATACCATGTCC 1560

Qy 1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC 1620
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC 1620

Qy 1621 CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCTGG 1680
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 CAGGCCTGTCACTCGTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCTGG 1680

Qy 1681 TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA 1740
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA 1740

Qy 1741 TGTGCCATCGTACTGCACTATGGGACTGGAACGCAGTGTGGAACGACACTCGCAACACTTA 1800
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 TGTGCCATCGTACTGCACTATGGGACTGGAACGCAGTGTGGAACGACACTCGCAACACTTA 1800

Qy 1801 CGAAGCATGCGGTGTAACAGCATGGCTAACACCGCACAACCGCATGGCAACAACTTAGCCCTGA 1860
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 CGAAGCATGCGGTGTAACAGCATGGCTAACACCGCACAACCGCATGGCAACAACTTAGCCCTGA 1860

Qy 1861 ATTGGCTATATCAATACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAAATTGGATGTC 1920
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 ATTGGCTATATCAATACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAAATTGGATGTC 1920

Qy 1921 AGGCCATTTGTATTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA 1980
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921 AGGCCATTTGTATTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA 1980

Qy 1981 TTCCACTCTCCTACCACGGAGAGGTGGCTAGGTTGCCCGGTTACCCCCACCTGTGGTACG 2040
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981 TTCCACTCTCCTACCACGGAGAGGTGGCTAGGTTGCCCGGTTACCCCCACCTGTGGTACG 2040

Qy 2041 TGGTTCTTGGTTACAGGTTTTACACGGTTTACAGTGAAAGACCTAGCCACAGG 2100
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Left column (positions 2041–3180):

```
Db  2041  TGGTTCTTGGTTACAGGTTCCGCGAAGGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG  2100
Qy  2101  ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGACCGGGTGC     2160
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  2101  ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGACCGGGTGC     2160
Qy  2161  TTTGTCTTACGGGAGTACCACCAGGCCGTGGTGCTAATTCTGTTGGGGTTGTGTGG       2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TTTGTCTTACGGGAGTACCACCAGGCCGTGGTGCTAATTCTGTTGGGGTTGTGTGG       2220
Qy  2221  CAGCAAGTATCTTATTTTAGCCTCGTGTACTTGTCCCCTTGTTTGGGCGCGCGCTTC     2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  CAGCAAGTATCTTATTTTAGCCTCGTGTACTTGTCCCCTTGTTTGGGCGCGCGCTTC     2280
Qy  2281  TGGTTACCCCTTTGCGTCCTGTCCCCATCCGTATCTCCCAGCTGGCTGGGATGT         2340
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db  2281  TGGTTACCCCTTTGCGTCCTGTCCCCATCCGTATCTCCCAGCTCGGCTGGGATGT        2340
Qy  2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTGCTTTGATTTTCTTCATCTGTCTATCTCCG      2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTGCTTTGATTTTCATCTGTCTGTCTATCTCCG    2400
Qy  2401  CTGCAGGCTACGTTATGCTGCCCCTTTTAGGGTTTGTGTCCCATGGCTGCGCGGCTTC    2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  CTGCAGGCTACGTTATGCTGCCCCTTTTAGGGTTTGTGTCCCATGGCTGCGCGGCTTC    2460
Qy  2461  AACTTCTTGTTGCGCAGCAGCTGCTGCCCAACCAGATTATGACTGGTGGGTGCGACTGCT  2520
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db  2461  AACTTCTTTGTTGCGCAGCAGCTGCTGCCCAACCAGATTACTGGTGGGTGCGACTGCT    2520
Qy  2521  AGTGGCAGGGTAGTTTGTGGGGCCGGCGTAACCGTGCTGCTCACCGCGATAGCTCTGCTTGT  2580
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  2521  AGTGGCAGGGTAGTTTGTGGGGCCGGCCGTGACCGTGCTGCTCACGGTAGCTCTGCTTGT  2580
Qy  2581  AGGTCCTTGGCCTCTCGGTAGCGCTTTTAACCCTCTTGGTTACGCGTTGGTTACGCCTGCTTCAGC  2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  AGGTCCTTGGCCTCTCGGTAGCGCTTTTAACCCTCTTGGTTACACGTACGCCTGCTTCAGC  2640
Qy  2641  TTTGATGACGAGATAATTGGAGGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT    2700
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  TTTGACACCGAGATAATTGGAGGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT   2700
Qy  2701  GTCTCGTTTGGCTTCTTGCTCACTTGTTGCTGTGCTTAGTTAACTCCTATCT           2760
          ||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db  2701  GTCTCGTTTTGGCTTCTTGCTCACTTGTTGCTGTGCTTAGTTAACTCCTATCT         2760
Qy  2761  TTTGGCAACGTTGGGAGAATTGGTTGTTTTGGAACGTAACAACTAACGGGAGGTTTTTCCT  2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  TTGGCAACGTTGGGAGAATTGGTTTGTTTTGGAACGTAACAACTAACGGGAGGTTTCTCCT  2820
Qy  2821  TGTGTCTGGTTTGTTGTTTCCCCGGTGCGACATATGACGCGCTGGTGACTTTCTGTGTGTCA  2880
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  2821  TGTGTCTGGTTTGTTGTTTCCCCGGTGCGACATATGACACGCTGGTGACTTTCTGTGTGTCA  2880
Qy  2881  CGTAGCTCTTCTTGCTTAACATCCAGTGCAGCATCGTTCTTTGGGACTGACTCTAGGGT     2940
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db  2881  CGTAGCTCTTCTTGCTTAACATCCAGTGCAGCATCGTTCTTTGGGACTGACTCTAGGGT    2940
Qy  2941  TAGGGCCCATAGAAATGTGGTGTTTTGGTGAGAAATGTGTGTTTTCATAAGCACTTGCA    3000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941  TAGGGCCCATAGAAATGTGGTGTTTTGGTGAGAAATGTGTGTTTTCATAAGCACTTGCA    3000
Qy  3001  TCTTAAGTTTTCCTTCTCGGGTTTTGGTTGGGTGTGTTTTTCTATAAGAGCCATTATGT   3060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  TCTTAAGTTTTCCTCTCGGGTTTTGGTGTGTTTTTTCTATAAGAGCCATTTTCCC       3060
Qy  3061  TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCAAGAGCACTTGCA      3120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061  TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCAAGAGCACTTGCA      3120
Qy  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACGCTTGGCGTGTGGGGACAC   3180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACACGTTGGCGTGTGGGGACAC  3180
```

Right column (positions 3181–4260):

```
Qy  3181  GGTTGATGGTTGCCCGGTTGTTGCGCGGTCTCGGCGCACCTTGTTTTCGCAGGGGTTGGCTAT  3240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181  GGTTGATGGTTGCCCGGTTGTTGCGCGGTCTCGGCGCACCTTGTTTTCGCAGGGGTTAGCTAT  3240
Qy  3241  GCCGCCAGATGGGTGGGCCATTACCGCACCTTTTACGCTGCAGTGTCTCTGAACGTGG      3300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241  GCCGCCAGATGGGTGGGCCATTACCGCACCTTTTACGCTGCAGTGTCTCTGAACGTGG     3300
Qy  3301  CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC    3360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301  CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC   3360
Qy  3361  TATCTTCAGATTAGGATCTCTGCCACTAGCTACATGGGATTTGTTTGTGACAACGTGTT    3420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361  TATCTTCAGATTAGGATCTCTGCCACTAGCTACATGGGATTTGTTTGTGACAACGTGTT   3420
Qy  3421  GTATACTGCTCACCATGGCCAAGGGGCGCCCGGTTGGCTCATCCCACGGCTCTCATACA    3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421  GTATACTGCTCACCATGGCCAAGGGGCGCCCGGTTGGCTCATCCCACAGGCTCCATACA   3480
Qy  3481  CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG   3540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481  CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG  3540
Qy  3541  GTCCCTTACTCGGTGCTCTTGCGGGGAGACCAAGGGGTATCTGTAACGACTGGGGGTC    3600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541  GTCCCTTACTCGGTGCTCTTGCGGGGAGACCAAGGGGTATCTGTAACGACTGGGGGTC   3600
Qy  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCCTATTGGTGTGCGTGCCCTCGGGCATGTTATTGG  3660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601  ATTGGTTGAGGTCAACAAATCCGATACCCCTATTGGTGTGCGTGCCCTCGGGCATGTTATTGG  3660
Qy  3661  GGCTGTTGCCAAGGGTCTTCAGGTGCCCGATTCTGTGCTCTCGGGCATGTGCAGCGCAAGTCAAC  3720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661  GGCTGTTGCCAAGGGTCTTCAGGTGCCCGATTCTGTGCTCTCGGGCATGTGCAGCGCAAGTCAAC  3720
Qy  3721  GATGTTCACCGCTGCTAGCAAGTTCTGGGGTTCAGTCAGTCAGATTCAGGGTTAGGCCGTT     3780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721  GATGTTCACCGCTGCTAGCAAGTTCTGGGGTTCAGTCAGTCAGATTCAGGGTTAGGCCGTT    3780
Qy  3781  GGTGTGTCTGGATACCATCCCCAGTACACAGCAGCACATGCCACTCTTGATACAAAACCTAC   3840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781  GGTGTGTCTGGATACCATCCCCAGTACACAGCAGCACATGCCACTCTTGATACAAAACCTAC  3840
Qy  3841  TGTGCCTAACGAGTATTCAGTGCAAATTTTAATTGCCCCCCACTGGCAGCGGCAAGTCAAC   3900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841  TGTGCCTAACGAGTATTCAGTGCAAATTTTAATTGCCCCCCACTGGCAGCGGCAAGTCAAC  3900
Qy  3901  CAAATTACCACTTTCTTACACATGCGAGGAGAAGTAATGAGGAGTATGAGGTCTTGGTCCTAAATCCCAGTGT  3960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901  CAAATTACCACTTTCTTCTTACACATGCGAGGAGAAGTAATGAGGAGTATGAGGTCTTGGTCCTAAATCCCAGTGT  3960
Qy  3961  GGCTACAACAGCATCAATGCCAAAGTACACGCCGGTCTTCACTTACGTACAGCACATATGGCAT  4020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961  GGCTACAACAGCATCAATGCCAAAGTACACGCCGGTCTTCACTTACGTACAGCACACATATGGCAT  4020
Qy  4021  CTATTTTAATGGCAAATGTACCAACACAGTGAATCATTTGTGACGAATGCATATCAATATGGCAT  4080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021  CTATTTTAATGGCAAATGTACCAACACAGTGAATCATTTGTGACGAATGCATATCAATATGGCAT  4080
Qy  4081  GTACCTGACCGGAGCAGCATGTTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGC  4140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081  GTACCTGACCGGAGCAGCATGTTCCCGGAACTATGATGTCATCATTGTGACGAATGCCATGC  4140
Qy  4141  TACCGATGCAACCAACCACCGTGTTCTTGGGCCATTGGAAAGGTTCAACCGAAGCTCCATCCAAAAA  4200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141  TACCGATGCAACCAACCACCGTGTTCTTGGGCCATTGGAAAGGTTCAACCGAAGTTCCATCCAAAAA  4200
Qy  4201  TGTTAGGCTAGGTCAGTGGTTCTTGTTCCACGCTCACCCCCCTGGAGTAATCCCTACCACACATGC  4260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201  TGTTAGGCTAGGTCAGTGGTTCTTGTTCCACGCTCACCCCCCTGGAGTAATCCCTACCACACATGC  4260
```

Qy 4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGTGAAAAAAGAT 4320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGTGAAAAAAGAT 4320

Qy 4321 TAAGGAGGAAAATCTGAAGAAAGGGAGACACCTTATCTTTGAGGCTACCAAAAAACACTG 4380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4321 TAAGGAGGAAAATCTGAAGAAAGGGAGACACCTTATCTTTGAGGCTACCAAAAAACACTG 4380

Qy 4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTTACTATAGGGG 4440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTTACTATAGGGG 4440

Qy 4441 ATGTGACATCTCAAAAATCCTGAGGGCGACTGTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4441 ATGTGACATCTCAAAAATCCCGAGGGCGACTGTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500

Qy 4501 TACAGGGTACACTGGTGACTTGATTCCGTGTATGACTCAGCCTCATGTGTAGAAGGCAC 4560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTCAGCCTCATGTGTAGAAGGCAC 4560

Qy 4561 ATGCCATGTTGACCTACCCTACTTTCACCATGGGTGTCGTGTGCGGGTTTCAGC 4620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4561 ATGCCATGTTGACCTACCCTACTTTCACCATGGGTGTCGTGTGCGGGTCTCTTCAGC 4620

Qy 4621 AATAGTTAAAGGCCAGCGTAGGGGCCGCACCAGGCCCGTGGGAGAGCTGGCATATACTACTA 4680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4621 AATAGTTAAAGGCCAGCGTAGGGGCCGCACCAGGCCCGTGGGAGAGCTGGCATATACTACTA 4680

Qy 4681 TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGTCAACATTGTTGAAGCCTT 4740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4681 TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGTCAACATTGTTGAAGCCTT 4740

Qy 4741 CGACGCAGCCAAGGCATGGTATGGTTTGTCATCAACAGAAGCTCATACTATTCTGACAC 4800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4741 CGACGCAGCCAAGGCATGGTATGGTTTGTCATCAACAGAAGCTCAAACTATTCTGACAC 4800

Qy 4801 CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTGGACGAGTGGGCTGATCT 4860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4801 CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTGGACGAGTGGGCTGATCT 4860

Qy 4861 CTTTTCTATGGTCAACCCCGAACCTTCATTTGTTCCTGAAAAAGAACTGCTGACAA 4920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4861 CTTTTCTATGGTCAACCCCGAACCTTCATTTGTTCCTGAAAAAGAACTGCTGACAA 4920

Qy 4921 TTATGTTTGTTGACTGCAGCCAACTACAACTGTGTCATCAGTATGGCTATGTGCTGCTCC 4980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4921 TTATGTTTGTTGACTGCAGCCAACTACAACTGTGTCATCAGTATGGCTATGTGCTGCTCC 4980

Qy 4981 CAATGACCGCACCGGTGGCAGGGAGCCCGGCTTGTGGGGTTCTGTG 5040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4981 CAATGACCGCACCGGTGGCAGGGAGCCCGGCTTGTGGGGTTCTGTG 5040

Qy 5041 GCGCTTGGACGCGCGCTGGCCTGTCTGGCCCCAGCGACTGACCAGATACCA 5100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5041 GCGCTTGGACGCGCGCTGGCCTGTCTGGCCCCAGCGACTGACCAGATACCA 5100

Qy 5101 AATGTGCTTCACTGAGTCAATAGCCATTTTGGCGCACTTTGTGTCGCGGGTTGCGGAGT 5160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5101 AATGTGCTTCACTGAGTCAATAGCCATTTTGGCGCACTTTGTGTCGCGGGTTGCGGAGT 5160

Qy 5161 GGCTATGGCTTATCTAGCCATTGACACTTTGACACTTTGGCGCACTTTGGCGCGGTTGCTCGGTC 5220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5161 GGCTATGGCTTATCTAGCCATTGACACTTTGACACTTTGGCGCACTTTGGCGCGGTTGCTCGGTC 5220

Qy 5221 TATTACATCAGTCCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAGAAATCGT 5280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5221 TATTACATCAGTCCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAGAAATCGT 5280

Qy 5281 GGAGGAGTGTCGTCATCATTCATTCCCCTTGGAGGCCATGGCCATGTTGCTGCAATTGACAAGCTGAA 5340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5281 GGAGGAGTGTCGTCATCATTCATTCCCCTTGGAGGCCATGGCCATGTTGCTGCAATTGACAAGCTGAA 5340

Qy 5341 GAGTACAATAACCACCACAACTAGTCCTTTCACATTGGAAACCGGCCCTTGAAAAACTTAACAC 5400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5341 GAGTACAATAACCACCACAACTAGTCCTTTCACATTGGAAACCGGCCCTTGAAAAACTTAACAC 5400

Qy 5401 CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGGTTTAGT 5460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5401 CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGGTTTAGT 5460

Qy 5461 CACTTTACCTGACAATCCCTTGCATCATGCGGTGTTGCTTCATTGCGGGTATTACTAC 5520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5461 CACTTTACCTGACAATCCCTTGCATCATGCGGTGTTGCTTCATTGCGGGTATTACTAC 5520

Qy 5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTGGAGGCGCAATTGCGTCCAA 5580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5521 CCCACTACCTCACAAGATCAAAAATGTTCCTGTCATTATTGGAGGCGCAATTGCGTCCAA 5580

Qy 5581 GCTTACAGACACGCTAGAGGCCGCACTGGCGTTCATGATGGCCGGGGCTGCGGGAACAGCTCT 5640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5581 GCTTACAGACGCGCTAGAGGCCGCACTGGCGTTCATGATGGCCGGGGCTGCGGGAACAGCTCT 5640

Qy 5641 TGGTACATGGACATCGGTGGGTTTGTCTTGACATGCTTCTTTGACATGCTAGGCGGCGTAGGCGGCTGCCGCCTC 5700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5641 TGGTACATGGACATCGGTGGGTTTGTCTTGACATGCTTCTTTGACATGCTAGGCGGCGTAGGCGGCTGCCGCCTC 5700

Qy 5701 ATCCACTGCTTGCTTGTGACATTTAAATGCTTGATGGGTGAGTGGCGCATATACTATGGATCAGCT 5760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5701 ATCCACTGCTTGCTTGTGACATTTAAATGCTTGATGGGTGAGTGGCGCATATACTATGGATCAGCT 5760

Qy 5761 TGCTGGTTAGTCTACTCCGCGTTCAATCCGGCCGCAGGAGTTGTGGGCGTCTTGTCAGC 5820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5761 TGCTGGTTAGTCTACTCCGCGTTCAATCCGGCCGCAGGAGTTGTGGGCGTCTTGTCAGC 5820

Qy 5821 TTGTGCAATGTTGCTTTGACAACAGCAGGGGCCAGATCACTGCCCAACAGACTTCTTAC 5880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5821 TTGTGCAATGTTGCTTTGACAACAGCAGGGGCCAGATCACTGCCCAACAGACTTCTTAC 5880

Qy 5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGACTACTTATTGCCCACTCGTGACACTCCG 5940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGACTACTTATTGCCCACTCGTGACATCCG 5940

Qy 5941 CAGGAAGATACTGGGCATTGGGAGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT 6000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5941 CAGGAAGATACTGGGCATTGGGAGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT 6000

Qy 6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGGCCTCATTGCTTGGGGTCTAGAGAT 6060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGGCCTCATTGCTTGGGGTCTAGAGAT 6060

Qy 6061 TTGGCAGTATGTGTGCAATTTCTTTGTGCAATTTTCTTTGTGTGATTTGCTTTAAATGTCCTTAAAGCTCCTTAAAGCTGGAGTTCA 6120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6061 TTGGCAGTATGTGTGCAATTTCTTTGTGCAATTTTCTTTGTGTGATTTGCTTTAAATGTCCTTAAAGCTCCTTAAAGCTGGAGTTCA 6120

Qy 6121 GAGCATGGTTAACATTCCTGGTGTCCTTTCACAGCTGCCAGAAGGGGCTGCCAGAAGGGGCCC 6180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6121 GAGCATGGTTAACATTCCTGGTGTCCTTTCACAGCTGCCAGAAGGGGCTGCCAGAAGGGGCCC 6180

Qy 6181 CTGGATTGGATCAGGTATGCTCCAAGCACGCTGTCCATGCGGTGCTGAACTCATCTTTTC 6240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6181 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGGCCTCATTGCTTGGGGTCTAGAGAT ... 6240

Qy 6241 TGTTGAGAATGGTTTTGCAAAACTTTACAAGGACCCCAGAACTTGTTCAAATTACTGGAG 6300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6241 TGTTGAGAATGGTTTTGCAAAACTTTACAAGGACCCCAGAACTTGTTCAAATTACTGGAG 6300

Qy 6301 AGGGGCTGTTCCAGTCAACGCTGCTACTGTCGCGTGTGTGGTCGGACCGACCCAACTGATTG 6360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6301 AGGGGCTGTTCCAGTCAACGCTGCTACTGTCGCGTGTGTGGTCGGACCGACCCAACTGATTG 6360

Qy 6361 GACTAGTTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAATGGGAGA 6420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6361 GACTAGTTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAATTGGGAGA 6420

Qy 6421 TCACATTTTGTTGTACAGCAGTATCCTCCCAAATGTCGTGTTCACCCAGGTGCCCCCAC 6480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTCACCCAGTGCCCCAAC 6480

Qy 6481 CTTGAGAGCTGCAGTGCCGTGGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGAGCCCAGT 6540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6481 CTTGAGAGCTGCAGTGCCGTGGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGAGCCCAGT 6540

Qy 6541 AACTCCTTGGACACATCTGCTTGCTTGCTCTGTCCGACGGGTAAAACTGTTAA 6600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6541 AACTCCTTGGACACATCTGCTTGCTTGCTCTGTCCGACGGGTAAAACTGTTAA 6600

Qy 6601 GCTTCCCTTCCGCGGTTGACGGTCACACACCTGGTGTCGCCATGCAACTAATTGCGTGA 6660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6601 GCTTCCCTTCCGCGGTTGACGGTCACACACCTGGTGTCGCCATGCAACTAATTGCGTGA 6660

Qy 6661 TGCACTTGAGACAAATGACTGTAATTCCATAAACAACACTCCTAGTGATGAAGCCGCAGT 6720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6661 TGCACTTGAGACAAATGACTGTAATTCCATAAACAACACTCCTAGTGATGAAGCCGCAGT 6720

Qy 6721 GTCCGGCTCCTGTTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGCAATTTC 6780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6721 GTCCGGCTCCTGTTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGCAATTTC 6780

Qy 6781 AGCTGGGCGTTGACACCACCACCAGTGCCAGCCCCTCCATCGAAGAGGTAGTGTAAGAAA 6840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6781 AGCTGGGCGTTGACACCACCACCAGTGCCAGCCCCTCCATCGAAGAGGTAGTGTAAGAAA 6840

Qy 6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTGCCTCCCCTCCGAGATCCGTCCC 6900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTGCCTCCCCTCCGAGATCCGTCCC 6900

Qy 6901 AGGAGTGTCATGTCTGCAACGAACGAAGTGACCCGTTAGAAGGTCCTTCAAACCT 6960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6901 AGGAGTGTCATGTCTGCAACGAACGAAGTGACCCGTTAGAAGGTCCTTCAAACCT 6960

Qy 6961 CCCTCCTTCACCACCTGTTCTACAGTTGCTGCAGTTGGCCGATGCCCCGTGTGGGGTGA 7020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6961 CCCTTCTTCACCACCTGTTCTACAGTTGCTGCAGTTGGCCGATGCCCCGTGTGGGGTGA 7020

Qy 7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCCAAACAGGCCGGAGGCCCTGATGA 7080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCCAAACAGGCCGGAGGCCCTGATGA 7080

Qy 7081 TTTACCCAGTTACCCTCCCAAAAAGGAGTCTCTGAATGGTCAGACGAAAGTTGGTCGAC 7140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db 7081 TTTACCCAGTTACCCTCCCAAAAAGGAGTCTCTGAATGGTCAGACGAAAGTTGGTCAAC 7140

Qy 7141 GGCTACAACCGCTTCCAGCTACTGTTCTACAGTTGGCCTCACTAGTGACCCGTACCCGGGAAAGGA 7200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7141 GACTACAACCGCTTCCAGCTACTGTTCTACAGTTGGCCTCACTAGTGACCCGTACCCGGGAAAGGA 7200

Qy 7201 TTCCACTCAGTCCAGCATGAGCTGGGGCATGGATCGGGCAAAACGGCCGCCAAAAAAGAGAGTT 7260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7201 TTCCACTCAGTCCAGCATGAGCTGGGGCATGGATCGGGCAAAACGGCCGCCAAAAAAGAGAGTT 7260

Qy 7261 TTCGTGCAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7261 TTCGTGCAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320

Qy 7321 TCTGTCTGCAACTCGGGCCATCATCGGGGCCATCTAGTGGTTTCCTCAAACAAAGATCATTGGTATGT 7380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7321 TCTGTCTGCAACTCGGGCCATCATCGGGGCCATCTAGTGGTTTCCTCAAACAAAGATCATTGGTATGT 7380

Qy 7381 GACTGAGCGCCGCGGGGATCGGGGACTTAGAAAGCAAAAAACAAGCTTCACTATTAATAGACAACCTCT 7440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7381 GACTGAGCGCCGCGGGGATCGGGGACTTAGAAAGCAAAAAACAAGCTTCACTATTAATAGACAACCTCT 7440

Qy 7441 GTTCCCCATCATACCACCACAAGCAAGTTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7441 GTTCCCCATCATACCACCACAAGCAAGTTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500

Qy 7501 CGGTGTCATGTGGGACTATGTGAATGAAGTAGCAGCTAGTGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 7560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7501 CGGTGTCATGTGGGACTATGTGAATGAAGTAGCAGCTAGTGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCTAAGTC 7560

Qy 7561 CCACATCACTGGCCTTCGGGGCACTGATGTTCGTTCGTGGAGCAGCCCGCAAGGCTGTTCT 7620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7561 CCACATCACTGGGCCTTCGGGGCACTGATGTTCGTTCGTGGAGCAGCCCGCAAGGCTGTTCT 7620

Qy 7621 GGACTTGCAGAAGTGTGCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7621 GGACTTGCAGAAGTGTGCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680

Qy 7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCAGAAACCAACAAAGAAACCCCAAG 7740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCAGAAACCAACAAAGAAACCCCAAG 7740

Qy 7741 GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGT 7800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7741 GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGT 7800

Qy 7801 TGCTCCTGACGTAGTTAAAGCTTCATGGGATGTCATGGGAGTGGTTTGTAGATCCACGTAC 7860
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7801 TGCTCCTGACGTAGTTAAAGCTTCATGGGATGTCATGGGAGTGGTTTGTGACCCACGTAC 7860

Qy 7861 CCGTGTCAAGCGTCTGTGTTGTCGATGTGGTCACCCGAGCGCACATGCGATAC 7920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7861 CCGTGTCAAGCGTCTGTGTTGTCGATGTGGTCACCCGAGCGCACATGCGATAC 7920

Qy 7921 AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTGGAGACAGACATCTACTC 7980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7921 AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTGGAGACAGACATCTACTC 7980

Qy 7981 AGCAGCTAAACTCAGTGACCAGCCTGCCATTCACACCACCAGGAGGCCAGTTATA 8040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 7981 AGCAGCTAAACTCAGTGACCAGCCTGCCATTCACACCACCAGGAGGCCAGTTATA 8040

Qy 8041 CGCTGGAGACCGATGATGCTTATGATGGCCGAGAGATCGGAGATGTGTAGGTC 8100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8041 CGCTGGGAGACCGATGATGCTTATGATGGCCGAGAGATCGGAGATGTGTAGGTC 8100

Qy 8101 TTCCGGCGTCTATACTACCTCAAGTTCCAACAGTTTGACCTGCTGGCTGAAGGTAAATGC 8160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8101 TTCCGGCGTCTATACTACCTCAAGTTCCAACAGTTTGACCTGCTGGCTGAAGGTAAATGC 8160

Qy 8161 TGCAGCCGAACAGGCTGGCATGAAGAACCTCGCTTCCTTATTGCGCGGCGATGATTGCAC 8220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8161 TGCAGCCGAACAGGCTGGCATGAAGAACCTCGCTTCCTTATTGCGCGGCGATGATTGCAC 8220

Qy 8221 CGTAATTTGGAAGAGCGCCGGAGCAGCAATGCAGACAAGAACAATGGCTGTGTCTTTGCTAG 8280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8221 CGTAATTTGGAAGAGCGCCGGAGCAGCAATGCAGACAAGAACAATGGCTGTGTCTTTGCTAG 8280

Qy 8281 CTGGATGAAGGTGATGGGTGCACCACAAGATTGTGTGCGCAACCCAAATACAGTTTGGA 8340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8281 CTGGATGAAGGTGATGGGTGCACCACAAGATTGTGTGCGCAACCCAAATACAGTTTGGA 8340

Qy 8341 AGAATTAACATCATGTCATCAATGTTACCTCGGCAGAACATGTAGTGGGCAAGCCTTA 8400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8341 AGAATTAACATCATGTCATCAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400

Qy 8401 CTACTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGTGCTCTCGCGAGGGTCTGGG 8460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8401 CTACTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGTGCTCTCGCGAGGGTCTGGG 8460

Qy 8461 ATACAACCCCAGTGCTGCGTGGTGTATCTAATACATCACTACCATGTTTGTGGGT 8520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8461 ATACAACCCCAGTGCTGCGTGGTGTATCTAATACATCACTACCATGTTTGTGGGT 8520

Qy 8521 TAGCCGTGTGTTGGCGTGTCCATTTCATGGAGGACAAAACTTCCCGA 8580
         ||||||||||||||||||||||||||||||||||||||||||||||
Db 8521 TAGCCGTGTGTTGGCGTGTCCATTTCATGGAGGACAAAACTTCCCGA 8580

Qy 8581 GACTGAGACCTTTGACTGGTGATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 8581 GACTGAGACCTTTGACTGGTGATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640

```
            FILING DATE: <Unknown>
     ATTORNEY/AGENT INFORMATION:
            NAME: POREMSKI, PRISCILLA E.
            REGISTRATION NUMBER: 33,207
            REFERENCE/DOCKET NUMBER: 5527.PC.01
     TELECOMMUNICATION INFORMATION:
            TELEPHONE: 708-937-6365
            TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 393:
     SEQUENCE CHARACTERISTICS:
            LENGTH: 9143 base pairs
            TYPE: nucleic acid
            STRANDEDNESS: double
            TOPOLOGY: linear
     MOLECULE TYPE: DNA (genomic)
     FEATURE:
            NAME/KEY: 5'UTR
            LOCATION: 1..445
     FEATURE:
            NAME/KEY: CDS
            LOCATION: 446..9037
     FEATURE:
            NAME/KEY: 3'UTR
            LOCATION: 9038..9143
     SEQUENCE DESCRIPTION: SEQ ID NO: 393:

US-08-467-344A-393

  Query Match     96.4%;  Score 9059.8;  DB 4;  Length 9143;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 9096;  Conservative 0;  Mismatches 37;  Indels 4;  Gaps 1;

Qy      1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG 60

Qy     61 CAGGGGCTGGGGCGTCGGGGATTTCCCCGCCCTCTGCGAGAAGGTGGAGCCAACCACCTTAGTAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAGGGGCTGGGGCGTCGGGGATTTCCCCGCCCTCTGCGAGAAGGTGGAGCCAACCACCTTAGTAT 120

Qy    121 GTAGGCGGCGGGACTCATGAGCCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTAGGCGGCGGGACTCATGAGCCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC 180

Qy    181 CCTGATGGGCGTTCATGGGTTCGGTGGTGGTGGCGCTTAGGCAGCCTCCAGCGCCCACCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCTGATGGGCGTTCATGGGTTCGGTGGTGGTGGCGCTTAGGCAGCCTCCAGCGCCCACCA 240

Qy    241 CCTCCCAGATAGAGCGGCGGCACTGTAGGGAAGACCCGGGAAGTACTACCAAGGACG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTCCCAGATAGAGCGGCGGCACTGTAGGGAAGACCCGGGAAGTACTACCAAGGACG 300

Qy    301 CAGACCTCTTTTGAGTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACCTACCACCTATATGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGACCTCTTTTGAGTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACCTACCACCTATATGT 360

Qy    361 TGGGATGGTTGGGGTTGGGGTTAGCCATACCGTACTGCCTGATAGGGTCCTTGGAGGGAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TGGGATGGTTGGGGTTGGGGTTAGCCATACCGTACTGCCTGATAGGGTCCTTGGAGGGAT 420

Qy    421 CTGGGAGTCTCTGTAGACCGTAGCACATGCCACTGTTATTCTACTCAAACAAGTCCTGTACC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGGGAGTCTCTGTAGACCGTAGCACATGCCACTGTTATTCTACTCAAACAAGTCCTGTACC 480

Qy    481 TGCGCCCAGAACGCGCAAGAACAACAAGACAGACCAGGCTTCATATCCTGTGTCCATTAAAAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGCGCCCAGAACGCGCAAGAACAACAAGACAGACCAGGCTTCATATCCTGTGTCCATTAAAAC 540

Qy    541 ATCTGTTGAAAGGGACAAGCGACAAGCGCAAAGCGCAAAGTCCAGCCGCGATGCTCGGCCTCGTAA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ATCTGTTGAAAGGGACAAGCGACAAGCGCAAAGCGCAAAGTCCAGCCGCGATGCTCGGCCTCGTAA 600

Qy    601 TTACAAAATTGGTATCCATGATGGCTTGCGACATTGGCTCAGGCTGGCTGCTTTGCCAGC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   8641 CATCATTGCTGGTGTGCACGGTATTGAGGCTTTCTCGGTGGTGCCTACACCAAGCGCTGA 8700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8641 CATCATTGCTGGTGTGCACGGTATTGAGGCTTTCTCGGTGGTGCCTACACCAAGCGCTGA 8700

Qy   8701 GATCCTCAGAGTTCCCAATCACTACAGACATGACATGACCCCCCTGGAGCCTGGCG 8760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8701 GATCCTCAGAGTTCCCAATCACTACAGACATGACATGACCCCCCTGGAGCCTGGCG 8760

Qy   8761 AAAGAAGCCAGGGCCGTCCTCGCCAGCGCCAAGAGAGCGTGGCGGAGCACACGCAAAATT 8820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8761 AAAGAAGCCAGGGCCGTCCTCGCCAGCGCCAAGAGAGCGTGGCGGAGCACACGCAAAATT 8820

Qy   8821 GGCTGCTTCCTTCCTCTCGGCATGCTACATCTAGACCTTCTACCAGATTTGGATAAGACGAG 8880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8821 GGCTGCTTCCTTCCTCTCGGCATGCTACATCTAGACCTTCTACCAGATTTGGATAAGACGAG 8880

Qy   8881 CGTGGCTCGGTACACACTTCAATTATTGTGATGTTTACTCCCCGGAGGGGGATGTGTT 8940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8881 CGTGGCTCGGTACACACTTCAATTATTGTGATGTTTACTCCCCGGAGGGGGATGTGTT 8940

Qy   8941 TATTACACCAGAGAAGAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTTT 9000
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   8941 TGTTACACCAGAGAAGAGATTGCAGAAGTTTCTGTGAAGTATTTGGCTGTCATTGTTTT 9000

Qy   9001 TGCCCTAGGGGCTCATTGCTGTTGACTAGCCATCAGCTAGCCCCAAATTCAAAATTAA 9060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9001 TGCCCTAGGGGCTCATTGCTGTTGACTAGCCATCAGCTAGCCCCAAATTCAAAATTAA 9060

Qy   9061 CTAACAG-----TTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC 9116
          |||||||     ||||||||||||||||||||||||||||||||||||||||||||||||
Db   9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC 9120

Qy   9117 GGGCTTAACGACCCCGC 9133
          |||||||||||||||||
Db   9121 GGGCTTAACGACCCCGC 9137
```

```
RESULT 7
US-08-467-344A-393
; Sequence 393, Application US/08467344A
; Patent No. 6586568
GENERAL INFORMATION:
     APPLICANT: JOHN N. SIMONS
                TAMI J. PILOT-MATIAS
                GEORGE J. DAWSON
                GEORGE G. SCHLAUDER
                SURESH M. DESAI
                THOMAS P. LEARY
                ANTHONY SCOTT MUERHOFF
                JAMES C. ERKER
                SHERI L. BUIJK
                ISA K. MUSHAHWAR
     TITLE OF INVENTION: NON-A, NON-B. NON-C. NON-D, NON-E HEPATITIS
                         REAGENTS AND METHODS FOR THEIR USE
     NUMBER OF SEQUENCES: 716
     CORRESPONDENCE ADDRESS:
          ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
          STREET: 100 ABBOTT PARK ROAD
          CITY: ABBOTT PARK
          STATE: IL
          COUNTRY: USA
          ZIP: 60064-3500
     COMPUTER READABLE FORM:
          MEDIUM TYPE: Floppy disk
          COMPUTER: IBM PC compatible
          OPERATING SYSTEM: PC-DOS/MS-DOS
          SOFTWARE: PatentIn Release #1.0, Version #1.25
     CURRENT APPLICATION DATA:
          APPLICATION NUMBER: US/08/467,344A
          FILING DATE: 07-Jun-1995
          CLASSIFICATION: <Unknown>
     PRIOR APPLICATION DATA:
          APPLICATION NUMBER: 08/424,550
```

```
Db

Db     601  TTACAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGCGTGCTTGCCAGC  660

Qy     661  TCATGGTTGGGACGCCAAGACCCTCGCCATAAGTCTCGCAATCTTCGGA  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661  TCATGGTTGGGACGCCAAGACCCTCGCCATAAGTCTCGCAATCTTCGGA  720

Qy     721  TTACCCTTTGGGGTGGATTGGTGATGTTACACTCACACACCTTCGGCCGCTGGT  780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721  TTACCCTTTGGGGTGGATTGGTGATGTTACACTCACACACCTTCGGCCGCTGGT  780

Qy     781  GGCAGGAGCGGTCGTTGACCAGTCTGCCAGATAGTACGCTTGCTGGAGGATGGAGTCAA  840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781  GGCAGGAGCGGTCGTTGACCAGTCTGCCAGATAGTACGCTTGCTGGAGGATCGGAGTCAA  840

Qy     841  CTGGGCTACTCGGTGGTTCGGTGTCCACCTTTTGTCGGTATGTGTGTATCTTGGCCTG  900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841  CTGGGCTACTCGGTGGTTCGGTGTCCACCTTTTGTCGGTATGTCGTGTATCTTGGCCTG  900

Qy     901  TCCCTGTAGTGGGGCGGGTCACTGACACCCAGACACAAATACCACAATCCTGACCAATTG  960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901  TCCCTGTAGTGGGGCGGGTCACTGACACCCAGACACAAATACCACAATCCTGACCAATTG  960

Qy     961  CTGCCAGCGTAATCAGGTGCGGACGAGTGCTGGGGTTCCCGCCAATCTGCCTACACGAGCCTGGTTG  1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961  CTGCCAGCGTAATCAGGTGCGGGACGAGTGCTGGGGTTCCCGCCAATCTGCCTACACGAGCCTGGTTG  1020

Qy    1021  TGTGATCTGCGGGACGAGTGCGTGGGGTTCCCGCCAATCGTAATCTCACCCCTTCCAA  1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021  TGTGATCTGCGGGACGAGTGCGTGGGGTTCCCGCCAATCGTAATCTCACCCCTTCCAA  1080

Qy    1081  TTGGACTGGCACGGACTCCTTCTTGGCTGACCACCATTTGTTGTGTGTTAGTGGTGACTG  1140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081  TTGGACTGGCACGGACTCCTTCTTGGCTGACCATTTGTTGTGTGTTAGTGGTGACTG  1140

Qy    1141  GACCTGTGACGCCCTTGACATTGGTAGTTGTGTGCGTGTATTGGTCGTGTATTAGTCCGTGACTG  1200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141  GACCTGTGACGCCCTTGACATTGGTAGTTGTGTGCGTGTATTGGTCGTGTATTAGTCCGTGACTG  1200

Qy    1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTGTTACCTGGA  1260
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTGTTACCTGGA  1260

Qy    1261  AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGATGGCCCGGCAAGGT  1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261  AGTGCCCACTGGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGATGGCCCGGCAAGGT  1320

Qy    1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCGCCTGATCCTCGACTATGCCTGACTAT  1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCGCCTGATCCTCTACTATGCCTGACTAT  1380

Qy    1381  GTTTAGCAGTGTACACTACCCTGCGGGTTGGCCGCTTATGCTTTACATAGAAGCCGACCCAT  1440
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381  GTTTAGCAGTGTACACTACCCTGCGGGTTGGCCGCTTATGCTTTACATAGAAGCCGACCCAT  1440

Qy    1441  GTGTAGTCAGTTGCTCTCCTAGCGCTTATGCTTTACATAGAAGCCGACCCTCTGGAAACCCTAT  1500
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441  GTGTAGTCAGTTGCTCTCCTAGCGCTTATGCTTTACATAGAAGCCGACCCTCTGGAAACCCTAT  1500

Qy    1501  CAGGGTGCCCACTGAGTTTGCCTCGCCTTGATGATACCATGTCC  1560
            ||||||||||||||||||||||||||||||||||||||||||||
Db    1501  CAGGGTGCCCACTGAGTTTGCCTCGCCTTGATGATACCATGTCC  1560

Qy    1561  TTGCCACTCTTATTTGAGTGAGAATGTGTCAGTGTTACAGTCCAAAGTGTGAC  1620
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561  TTGCCACTCTTATTTGAGTGAGAATGTGTCAGTGTTACAGTCCAAAGTGTGGAC  1620

Qy    1621  CAGGCCTATCACTCTAGAGTATAACACTCCATATCTTGGTACCCCTATACAATCCCTGG  1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621  CAGGCCTATCACTCTAGAGTATAACACTCCATATCTTGGTACCCCTATACACATCCCTGG  1680

Qy    1681  TGCGAGGGGATGTATGGTTAAATTCAAAAATGTTAAATTCAAAATAACACATGGGGTTGCTGCCGTATTCGCAA  1740
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681  TGCGAGGGGATGTATGGTTAAATTCAAAAATGTTAAATTCAAAATAACACATGGGGTTGCTGCCGTATTCGCAA  1740
```

```
Qy

Db    1741  TGTGCCATCGTACTGCACTATGGGCACTGATGCAGTGTGAAGAGCACTCGCAACACTTA  1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1741  TGTGCCATCGTACTGCACTATGGGCACTGATGCAGTGTGGAACGACACTCGCAACACTTA  1800

Db    1801  CGAAGCAATGGCGGTGTAACACCATGGCTAACACCATGGCTAACAACGGCTCAGCGCCCTGAA  1860
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1801  CGAAGCAATGGCGGTGTAACACCATGGCTAACACCATGGCTAACAACGGCTCAGCCTGAA  1860

Db    1861  ATTGGCTATATTACAATACCCTGGGCTCAAAGAAATGTTTAAACCTCATAATTGGATGTC  1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1861  ATTGGCTATATTACAATACCCTGGGCTCAAAGAAATGTTTAAACCTCATAATTGGATGTC  1920

Db    1921  AGGCCATTGTATTTTGAGGGATCAGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAA  1980
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1921  AGGCCATTGTATTTTGAGGGATCAGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAA  1980

Db    1981  TTCCACTCTCCTACCACCGGAGAGGTGGGCTAGGTGCCGGTACCCACCTGTGGTACG  2040
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1981  TTCCACTCTCCTACCACCGGAGAGGTGGGCTAGGTGCCGGTACCCACCTGTGGTACG  2040

Db    2041  TGGTTCTTGGTTACAGGTTTACAGTGTGATGTGAAAGAACTAGCCACAGG  2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    2041  TGGTTCTTGGTTACAGGTTTACAGTGTGATGTGAAAGAACTAGCCACAGG  2100

Db    2101  ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC  2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    2101  ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC  2160

Db    2161  TTTGTCTCCTTACGGGAGTTACCACCAAGGCCGTGGTGGTCTAATTCGTTGGGGTTGTGTGG  2220
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    2161  TTTGTCTCCTTACGGGAGTTACCACCAAGGCCGTGGTGGTCTAATTCGTTGGGGTTGTGTGG  2220

Db    2221  CAGCAAGTATCTTATTTTAGCTACCTCTGTCCCATCCCAGTCGTATCTCCAGTCGTATCTCCAGTCGTATCCC  2280
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    2221  CAGCAAGTATCTTATTTTAGCTACCTCTGTCCCATCCCAGTCGTATCTCCAGTCGTATCTCCC  2280

Db    2281  TGGTTACCCTTTGCGTCCTGTGCTCGTATCCCAGTCGTATCTCCAAGTCGGCTGGGATGT  2340
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    2281  TGGTTACCCTTTGCGTCCTGTGCTCGTATCCCAGTCGTATCTCCAAGTCGGCTGGGATGT  2340

Db    2341  TTTGTCTAAAGCTCAAGTAGCTCCTTCTTCATCTCGTGATTTTCTTCATCTCGTGCTATCTCCG  2400
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    2341  TTTGTCTAAAGCTCAAGTAGCTCCTTCTTCATCTCGTGATTTTCTTCATCTCGTGCTATCTCCG  2400

Db    2401  CTGCAGGCTACGTTATGCTGCCATGGGTTGTGCCCATGGCTGCCTGCGGGGCGCCTGCCCCT  2460
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    2401  CTGCAGGCTACGTTATGCTGCCATGGGTTGTGCCCATGGCTGCCTGCGGGGCGCTTGCCCCT  2460

Db    2461  AACTTTCTTTGCTGCAGCAGCGCTGCTGCTGCCCAACCAGATTATGACTGGTGTGGGGCGCATAGCTGCTTGT  2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    2461  AACTTTCTTTGCTGCAGCAGCGCTGCTGCTGCCCAACCAGATTATGACTGGTGTGGGGCGCATAGCTGCTTGT  2520

Db    2521  AGTGGCAGGGTAGTTTTTGTGGGGCCGCGTAACCCTGACCGTGGTTCCAGTATAGCTCTGCTTGT  2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    2521  AGTGGCAGGGTAGTTTTTGTGGGGCCGCGTAACCCTGACCGTGGTTCCAGTATAGCTCTGCTTGT  2580

Db    2581  AGTCCTTGGCCTCTGGTAGCGCTTTAACCCTCGTTGGTTACCTGCTTCAGC  2640
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    2581  AGTCCTTGGCCTCTGGTAGCGCTTTAACCCTCGTTGGTTACCTGCTTCAGC  2640

Db    2641  TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGTTAGTTGTCAT  2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    2641  TTTTGACACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGTTAGTTGTCAT  2700

Db    2701  GTCTCGTTTGGCCTTCTTTGCTCACTTGTTACCTCGCTGGTTAACTCCTATCT  2760
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    2701  GTCTCGTTTGGCCTTCTTTGCTCACTTGTTACCTCGCTGTTAACTCCTATCT  2760

Db    2761  TTGGCAACGTTGGGGAATTGGTTTTGGAACGTTACACTGAAGACGGGTTTTCCTT  2820
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    2761  TTGGCAACGTTGGGGAATTGGTTTTGGAACGTTACACTGAAGACGGAGGTTTTCTCCTT  2820
```

```
Qy  2821  TGTGCTGGTTGTTCCCCGGTGCGTCCCCGGTGCGACATATGACCGCGTGGTGACTTCTGTGTGTGTCA  2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  TGTGCTGGTTGTTTTCCCCGGTGCGTCCCCGGTGCGACATATGACACGCGTGGTGACTTCTGTGTGTGTCA  2880

Qy  2881  CGTAGCTCTTCTCTATGTTTAACATCCAGTGCAGCATCGTTCTTGGGACTGACTCTAGGGT  2940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  CGTAGCTCTTCTCTATGTTTAACATCCAGTGCAGCATCGTTCTTGGGACTGACTCTAGGGT  2940

Qy  2941  TAGGGCCCATAGAAGTTGGTGCGTCTCGGAAAGTGTCATGTCTTGTATTCTCATTATGT  3000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941  TAGGGCCCATAGAAGTTGGTGCGTCTCGGAAAGTGTCATGTCTTGTATTCTCATTATGT  3000

Qy  3001  TCTTAAGTTTTCCTCCTTAGTGTTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA  3060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  TCTTAAGTTTTTCCTCCTTAGTGTTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA  3060

Qy  3061  TGGTGATGTCTTGCCTAATGATTTTGCCTCCGAAACTACCATTGCAAGAGCCATTTTCCC  3120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061  TGGTGATGTCTTGCCTAATGATTTTGCCTCCGAAACTACCATTGCAAGAGCCATTTTTCCC  3120

Qy  3121  TTTTGAAGGCAAGGGTCTATAGGAGAAGACGCTTGGCGTGTGGGACAC  3180
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  TTTTGAAGGCAAGGGTCTATAGGAGAAGACGCTTGGCGTGTGGGACAC  3180

Qy  3181  GGTTGATGGTTGCCCGTTGTGCCGTGTCTCGGCGACCTTGTTTTCGCAGGGTTGGCTAT  3240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181  GGTTGATGGTTGCCCGTTGTGCCGTGTCTCGGCGACCTTGTTTTCGCAGGGTTAGCTAT  3240

Qy  3241  GCCGCCAGATGGGTGGGCCATTACCCGCCACTTTTACGCTGCGAGTGTCTCTCTGAACTGG  3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241  GCCGCCAGATGGGTGGGCCATTACCCGCCACTTTTACGCTGCGAGTGTCTCTCTGAACGTGG  3300

Qy  3301  CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCGAACTTGGACTGGAAC  3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301  CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCGAACTTGGACTGGAAC  3360

Qy  3361  TATCTTCAGATTAGGATCTCTGGCCACTAGCTACACTGGGATTTGTTTGTGTGACAACGTGTT  3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361  TATCTTCAGATTAGGATCTCTGGCCACTAGCTACACTGGGATTTGTTTGTGTGACAACGTGTT  3420

Qy  3421  GTATACTGCTCACCATGGCAGCAGGAGGCGCCGGTTGGCTCATCCCACCAGGCTCTATACA  3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421  GTATACTGCTCACCATGGCAGCAGGAGGCGCCGGTTGGCTCATCCCACCAGGCTCCATACA  3480

Qy  3481  CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG  3540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481  CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG  3540

Qy  3541  GTCCCTTACTCGGTGCTCTTGCGGTGTATCTGGTAACAGACTGGGGTC  3600
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541  GTCCCTTACTCGGTGCTCTTGCGGTGTATCTGGTAACAGACTGGGGGTC  3600

Qy  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCCTTATTGGTGTGTGTGGCGCGGGGCCTTCCCAT  3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCCTTATTGGTGTGTGTGGCGGGGGCCCTTCCCAT  3660

Qy  3661  GGCTGTGTGCCAAGGGTTCTTGCCAGGTTGCCCGATTCGTGTGCGTCTTCCTCCGGGCATGTTATTGG  3720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661  GGCTGTGTGCCAAGGGTTCTTGCCAGGTTGCCCGATTCGTGTGCGTCTTCCTCCGGGCATGTTATTGG  3720

Qy  3721  GATGTTCACCCGTGCTAGAAATTCTGGCGTTCAGTCAGTCAGCCAGATTAGGGTTAGGCCGTT  3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721  GATGTTCACCCGTGCTAGAAATTCTGGCGTTCAGTCAGTCAGCCAGATTAGGGTTAGGCCGTT  3780

Qy  3781  GGTGTGTGCTGGATACCATCCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC  3840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781  GGTGTGTGCTGGATACCATCCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC  3840

Qy  3841  TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGCAGCGGCAAGTCAAC  3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841  TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGCAGCGGCAAGTCAAC  3900

Qy  3901  CAAATTACCACTTTCTTACATGCAGGAGAAGTATGATGCCTTGGTCTCCTAAATCCCAGTGT  3960
```

```
Qy  3901  CAAATTACCACTTTCTTACATGCAGGAGAAGTATGATGCCTTGGTCTCCTAAATCCCAGTGT  3960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961  CAAATTACCACTTTCTTACATGCAGGAGAAGTATGATGCCTTGGTCTCCTAAATCCCAGTGT  4020

Qy  3961  GGCTACACAGCATCAATGCCAAAGTACATGCAGGGGACGCGGACGTATGTGGTCTGAATCCAAATTG  4020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961  GGCTACACAGCATCAATGCCAAAGTACATGCAGGGGACGCGGACGTATGTGGTCTGAATCAAATTG  4020

Qy  4021  CTATTTTAATGGCAAATGCATCCAACACAGCAGGGGCTTCACTTACGACTACAGCACACATATGGCAT  4080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021  CTATTTTAATGGCAAATGCATCCAACACAGCAGGGGCTTCACTTACGACTACAGCACACATATGGCAT  4080

Qy  4081  GTACCTGACCGGAGCATGTCCCGGAACTATGATGTGAATCATTGTGACGAATGCCATGC  4140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081  GTACCTGACCGGAGCATGTCCCGGAACTATGATGTGAATCATCATTGTGACGAATGCCATGC  4140

Qy  4141  TACCGATGCAACCACCGTGTTGGGCATTGGGAAAGGTCCTAACCGAAGGTCCATCCAAAAA  4200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141  TACCGATGCAACCACCGTGTTGGGCATTGGGAAAGGTTCTAACCGAAGGTCCATCCAAAAA  4200

Qy  4201  TGTTAGGCTAGTTGGTTCTTGCCACGCGGCTACCCCCCCTGGAGTAATCCTACCACATGC  4260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201  TGTTAGGCTAGTTGGTTCTTGCCACGCGGCTACCCCCCCTGGAGTAATCCTACCACATGC  4260

Qy  4261  CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGGAAAAAGAT  4320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261  CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGGAAAAAGAT  4320

Qy  4321  TAAGGAGGAAAATCTGAAGAAAGGAGACGACACCTTATCTTTGAGGCTACCCAAAAAACACTG  4380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321  TAAGGAGGAAAATCTGAAGAAAGGAGACGACACCTTATCTTTGAGGCTACCCAAAAAACACTG  4380

Qy  4381  TGATGAGCTTGCTAACGAGTTAGTCTGAAAGGGAATAACAGTGTCTCTTACTATAGGGG  4440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381  TGATGAGCTTGCTAACGAGTTAGTCTGAAAGGGAATAACAGTGTCTCTTACTATAGGGG  4440

Qy  4441  ATGTGACATCTCAAAAATCCCTGAGGCGACTGTGTAGTAGTGCCACTGATGGCCTTGTG  4500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441  ATGTGACATCTCAAAAATCCCTGAGGCGACTGTGTAGTAGTGCCACTGATGGCCTTGTG  4500

Qy  4501  TACAGGGTACACTGGTGACTTTGATTCCGTGATGGGCTCATGGTAGAAGGCAC  4560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501  TACAGGGTACACTGGTGACTTTGATTCCGTGATGGGCTCATGGTAGAAGGCAC  4560

Qy  4561  ATGCCATGTTGACCCTACTTTCGTGTGTGTGCGGGTTCAGC  4620
          |||||||||||||||||||||||||||||||||||||||||
Db  4561  ATGCCATGTTGACCCTACTTTCGTGTGTGTGCGGGTTCTCAGC  4620

Qy  4621  AATAGTTAAAGGCCAGCGGCCCGCACAGCCGTGGGAGAGCTGGCATATACTACTA  4680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621  AATAGTTAAAGGCCAGCGGCCCGCACAGCCGTGGGAGAGCTGGCATATACTACTA  4680

Qy  4681  TGTAGACGGGAGTTGTACCCCTCGGGTATGGTGTGCGGGGTTCTTGGAAGCCTT  4740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681  TGTAGACGGGAGTTGTACCCCTCGGGTATGGTGTGCGGGGTTCTTGGAAGCCTT  4740

Qy  4741  CGACGCAGCCAAGGCCATGGTAGTATGGTTTGTCATCATCAACAGAGCAAGCTCAAACTACTATTGGACAC  4800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741  CGACGCAGCCAAGGCCATGGTAGTATGGTTTGTCATCATCAACAGAGCAAGCTCAAACTATTCTCGGACAC  4800

Qy  4801  CTATCGCACCCAACCTGGTTACCTGCGATTAGGGTTAGGCCGTT  4860
          ||||||||||||||||||||||||||||||||||||||||||||
Db  4801  CTATCGCACCCAACCTGGTTACCTGCGATAGGAGCAAATTTGGACGTGGGCTGATCT  4860

Qy  4861  CTTTTCATGTCAACCCGAACCTTCATTTGTCAATACTGCGAAAAGAACTGCTGACAA  4920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861  CTTTTCATGTCAACCCGAACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA  4920

Qy  4921  TTATGTTTGTTGACTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCC  4980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921  TTATGTTTGTTGACTGCAGCGCCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCC  4980

Qy  4981  CAATGACGCACCACGGTGGCAGGAGCCGGCGGTTGGAAAAACCTTGTGGGGTTCTGTG  5040
```

```
Db  4981  CAATGACGCACCACGGTGGCAGGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG  5040
                                                                         
Qy  5041  GCGCTTGGACGGGCGCTGACGCGGCCGTGTCCTGGCGCCTGCCAGAGCCCAGCGAGGTGACCAGATACCA  5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041  GCGCTTGGACGGGCGCTGACGCGGCCGTGTCCTGGCGCCTGCCAGAGCCCAGCGAGGTGACCAGATACCA  5100

Qy  5101  AATGTGCTTCACTGAAGTCAATACTTCTGGACAGCCCACTCGTCTGTGGGCGTTGGGAGT  5160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101  AATGTGCTTCACTGAAGTCAATACTTCTGGACAGCCCACTCGTCTGTGGGCGTTGGGAGT  5160

Qy  5161  GGCTATGGCTTATCTGACACTTTGGCGCCACTTTGGCGCCACTCGTGTGTGGGCGTTGCTGTC  5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161  GGCTATGGCTTATCTGACACTTTGGCGCCACTTTGGCGCCACTCGTGTGTGGGCGTTGCTGTC  5220

Qy  5221  TATTACATCAGTCCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGGTTGACGAAGAAATCGT  5280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221  TATTACATCAGTCCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGGTTGACGAAGAAATCGT  5280

Qy  5281  GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGGCTGCAATTGACAAGCTGAA  5340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281  GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGGCTGCAATCGATAAGCTGAA  5340

Qy  5341  GAGTACAATCACCACAACTAGTCCTTTCACATTGGAAACCGGCCCTTGAAAAACTTAACAC  5400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341  GAGTACAATAACCACAACTAGTCCTTTCACATTGGAAACCGGCCCTTGAAAAACTTAACAC  5400

Qy  5401  CTTTCTTGGGGCCTCATGGACCTACAATCCTTGCTATCATAGAGTATTGCTGTGGTTAGT  5460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401  CTTTCTTGGGGCCTCATGGACCTACAATCCTTGCTATCATAGAGTATTGCTGTGGTTAGT  5460

Qy  5461  CACTTTACCTGACAATCCCTTTGCATCATGCGTGTTTCATTGCTTTCATTGCGGGTATTACTAC  5520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461  CACTTTACCTGACAATCCCTTTGCATCATGCGTGTTTCATTGCTTTCATTGCGGGTATTACTAC  5520

Qy  5521  CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTGGAGGCCAATTGCGTCCAA  5580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521  CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTGGAGGCCAATTGCGTCCAA  5580

Qy  5581  GCTTACAGACGCTAGAGGCGCACTGGCGTTCATGATGGCCGGGCGCTGCGGGAACAGCTCT  5640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581  GCTTACAGACGCTAGAGGCGCACTGGCGTTCATGATGGCCGGGCGCTGCGGGAACAGCTCT  5640

Qy  5641  TGGTACATGGACATGGTGGGTTTGCTTTGACATGCTAGGCGGCATGCTGCCGCGCCTC  5700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641  TGGTACATGGACATGGTGGGTTTGCTTTGACATGCTAGGCGGCATGCTGCCGCGCCTC  5700

Qy  5701  ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGAGTCGCCCACTATGGATCAGCT  5760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701  ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGAGTCGCCCACTATGGATCAGCT  5760

Qy  5761  TGCTGGTTTAGTCTACTCCGGGCGTTCAATCCGGGCGGCCAGGAGTGTGTGGGCGTCTTGTCAGC  5820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761  TGCTGGTTTAGTCTACTCCGGGCGTTCAATCCGGGCGGCCAGGAGTGTGTGGGCGTCTTGTCAGC  5820

Qy  5821  TTGTGCAATGTTTGCTTTGACAACAGCAGGGCATCACTGCCCAACAGACTTCTTAC  5880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821  TTGTGCAATGTTTGCTTTGACAACAGCAGGGCATCACTGCCCAACAGACTTCTTAC  5880

Qy  5881  TATGGCTTGGCTCCACACCCGACGGAGATGATTGCGGCCTCATTGCCTACCACTCGTTGACATCCG  5940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881  TATGGCTTGGCTCCACACCCGACGGAGATGATTGCGGCCTCATTGCCTACCACTCGTTGACATCCG  5940

Qy  5941  CAGGAGATACTGGGCATGGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT  6000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941  CAGGAAGATACTGGGCATGGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT  6000

Qy  6001  CCGTTGGCTCCACACCCGACGGAGCATTCTCGGAGGCATCCCCCTGGAGTGTCTTAGAGAT  6060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001  CCGTTGGCTCCACACCCGACGGAGCATTCTCGGAGGCATCCCCCTGGAGTGTCTTAGAGAT  6060

Qy  6061  TTGGCAGTATGTGTGCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGTCTGGAGTTCA  6120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061  TTGGCAGTATGTGTGCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGTCTCGGAGTTCA  6120
```

```
Qy  6121  GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAGGGGTACAAGGGCCC  6180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121  GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAGGGGTACAAGGGCCC  6180

Qy  6181  CTGGATTGGATCAGGTATGCTCCAAGCACGCGTCCATGCGGTGTGAACTCATCTTTC  6240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181  CTGGATTGGATCAGGTATGCTCCAAGCACGCGTCCATGCGGTGTGAACTCATCTTTC  6240

Qy  6241  TGTTGAGAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241  TGTTGAGAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300

Qy  6301  AGGGGCTGTTCCAGTCAACGCTAGGCTGTGTGTGGTCGGCTAGACCGGACCCAACTGATTG  6360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301  AGGGGCTGTTCCAGTCAACGCTAGGCTGTGTGTGGTCGGCTAGACCGGACCCAACTGATTG  6360

Qy  6361  GACTAGTCTGTCGTCAATTATGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA  6420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361  GACTAGTCTGTCGTCAATTATGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA  6420

Qy  6421  TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGTGCCCCCAAC  6480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421  TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGTGCCCCCAAC  6480

Qy  6481  CTTGAGAGCTGCAGTGGCCGTGGACGGTTCAGTGTTATCTAGGTGAGCCCAA  6540
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481  CTTGAGAGCTGCAGTGGCCGTGGACGGTTCAGTGTTATCTAGGTGAGCCCAA  6540

Qy  6541  AACTCCTTGGACACAAATGACTGTAATTCCACAACAACACTCCTAGTGATGAAGCCGCAGT  6600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541  AACTCCTTGGACACAAATGACTGTAATTCCATAACAACACTCCTAGTGATGAAGCCGCAGT  6600

Qy  6601  GCTTCCCTTCCGCGGTTCACACCGGTGTGGCGCCATGCGGCATGCCAGCCCCTTCATTGCGTGA  6660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601  GCTTCCCTTCCGCGGTTCACACCGGTGTGGCGCCATGCGGCATGCCAGCCCCTTCAATTGCGTGA  6660

Qy  6661  TGCACTTGAGACAAATGACTGTAATTCCACAAACAACTCCTAGTGATGATGAAGCCGCAGT  6720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661  TGCACTTGAGACAAATGACTGTAATTCCATAACACTCCTAGTGATGATGAAGCCGCAGT  6720

Qy  6721  GTCCGCTCCTGTTCAAACAGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC  6780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721  GTCCGCTCCTGTTCAAACAGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC  6780

Qy  6781  AGTGGCGTTGACACCACCAAACTGCCAGCCCCTCATCGAAGAGTAGTGGTAAGAAA  6840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781  AGTGGCGTTGACACCACCAAACTGCCAGCCCCTCATCGAAGAGTAGTGGTAAGAAA  6840

Qy  6841  GCGCCAGTTCCGGGCAAGAACTGTTCGCTTACCTTGCCTCCCCCTCCCGAGATCCGTCCC  6900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841  GCGCCAGTTCCGGGCAAGAACTGTTCGCTTACCTTGCCTCCCCCTCCCGAGATCCGTCCC  6900

Qy  6901  AGGAGTGTCATGTCCTGAAAGCCTGCAACGAAGCTGACCCGTTGAGAAGGTCCTTCAAACCT  6960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901  AGGAGTGTCATGTCCTGAAAGCCTGCAACGAAGCTGACCCGTTGAAAGGTCCTTCAAACCT  6960

Qy  6961  CCCTTCCTTCCACACCTGTTCTACAGTTGCCACTCGTTGGGAGCCGGGTGA  7020
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961  CCCTTCCTTCCACACCTGTTCTACAGTTGCCACTCGTTGGGAGCCGGGTGA  7020

Qy  7021  GTGTAACCCTTCACTGACAATTGGATGTGCAATGACCGAAACAGGCGGAGGGCCCTGATGA  7080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021  GTGTAACCCTTCACTGACAATTGGATGTGCAATGACCGAAACAGGCGGAGGGCCCTGATGA  7080

Qy  7081  TTTACCCAGTTACCCTCCCCAAAAAGAAGGAGTCTCTGAAATGGTCAGCGGAAGTTGGTCAAC  7140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081  TTTACCCAGTTACCCTCCCCAAAAAGAAGGAGTCTCTGAAATGGTCAGCGGAAGTTGGTCAAC  7140

Qy  7141  GGCTACAACCGCTTCCAGCTACGTTACTCGCCCCGTACCCTAAGATACGGGAAAGGA  7200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141  GACTACAACCGCTTCCAGCTACTGGACGTTACTCGCCCGTACCCTAAGATACGGGAAAGGA  7200
```

```
Qy  7201 TTCCACTCAGTCAGCCCCGCCAAACGGCTACAAAAAGAAGTTGGGAAAGAGTGAGTT 7260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201 TTCCACTCAGTCAGCCCCGCCAAACGGCTACAAAAAGAAGTTGGGAAAGAGTGAGTT 7260

Qy  7261 TTCGTGCAGCATGAGTGAGCCGACGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261 TTCGTGCAGCATGAGTGAGCCGACGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320

Qy  7321 TCTGTGTCGCAACTCGGGCCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGT 7380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7321 TCTGTGTCGCAACTCGGGCCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGT 7380

Qy  7381 GACTGAGCCGCGGGATGCGGAGCTTAGAAACAAAAGTCACTATTAATAGACAACCTCT 7440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381 GACTGAGCCGCGGGATGCGGAGCTTAGAAACAAAAGTCACTATTAATAGACAACCTCT 7440

Qy  7441 GTTCCCCCATCATACCACCAACAAGCAAGTGAGATTGGCTAAGAAAAAGCTTCAAAGTTGT 7500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441 GTTCCCCCATCATACCACCAACAAGCAAGTGAGATTGGCTAAGAAAAAGCTTCAAAGTTGT 7500

Qy  7501 CGGTGTCATGTGGACTATGATGAAGTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTC 7560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7501 CGGTGTCATGTGGACTATGATGAAGTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTC 7560

Qy  7561 CCACATCACTGCCTTCGGGGCACTGATGTTCGTTCGTGGAGCAGCCCGCCAAGGCTGTTCT 7620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7561 CCACATCACTGCCTTCGGGGCACTGATGTTCGTTCGTGGAGCAGCCCGCCAAGGCTGTTCT 7620

Qy  7621 GGACTTGCAGAAGTGTGTCGAGGCAGTGAGTGAGATCATTATCGGCAAAACTGTGAT 7680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7621 GGACTTGCAGAAGTGTGTCGAGGCAGTGAGTGAGATCATTATCGGCAAAACTGTGAT 7680

Qy  7681 AGTTCCAAAGGAGGAGTCTTCGTGAGAAGAAACAACAAAGAAACCCCCAAG 7740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7681 AGTTCCAAAGGAGGAGTCTTCGTGAGAAGAAACAACAAAGAAACCCCCAAG 7740

Qy  7741 GCTTATCTCGTACCCCCACCTTGAAATGAGTGTTGAGAAGATGTTACTACGGGTCAGGT 7800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7741 GCTTATCTCGTACCCCCACCTTGAAATGAGTGTTGAGAAGATGTTACTACGGGTCAGGT 7800

Qy  7801 TGCTCCTGACGTAGTAGTAAAGCTGTCATGGGAGATGCGGTTTGTAGATCCACGGTAC 7860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7801 TGCTCCTGACGTAGTAGTAAAGCTGTCATGGGAGATGCGGTTTGTAGATCCACGGTAC 7860

Qy  7861 CCGGTGTCAAGGCTCTCGTGTCGTGTCACCCGATGCAGTCGGAGCCACATGCGGCGATAC 7920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7861 CCGGTGTCAAGGCTCTCGTGTCGTGTCACCCGATGCAGTCGGAGCCACATGCGGCGATAC 7920

Qy  7921 AGTGTGTTTGACAGTACCATCACACACCGGAGGATATCATGTGGCGACAGACAGATCCTCTC 7980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7921 AGTGTGTTTGACAGTACCATCACACACCGGAGGATATCATGTGGCGACAGACAGATCCTCTC 7980

Qy  7981 AGCAGCCTAAACTCAGTGACCAACACGTGGCTTATGATGGCCGAGATCGTAGGTGTTAGGTC 8040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7981 AGCAGCCTAAACTCAGTGACCAACACGTGGCTTATGATGGCCGAGATCGTAGGTGTTAGGTC 8040

Qy  8041 CGCTGAGGAGGACCGATCGCTTATGATGGGCCGAGATATCCGAGATATCGTAGGTGTTAGGTC 8100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8041 CGCTGAGGAGGACCGATCGCTTATGATGGGCCGAGATATCCGAGATATCGTAGGTGTTAGGTC 8100

Qy  8101 TTTCCGGCGTCTATACTCCAAGTTCAACAGTTGACCTGCTCTGGCTGAAGGTAAATGC 8160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8101 TTTCCGGCGTCTATACTCCAAGTTCAACAGTTGACCTGCTCTGGCTGAAGGTAAATGC 8160

Qy  8161 TGCAGCCGAACAGGCTGGCATGAAGAACCCTCGCTTCCTTCGCTTCCTCGCTTCCTCGCGGCGATGATTGCAC 8220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161 TGCAGCCGAACAGGCTGGCATGAAGAACCCTCGCTTCCTTCGCTTCCTCGCTTCCTCGCGGCGATGATTGCAC 8220

Qy  8221 CGTAATTGGAAGAGCCGCGGAGCAGAGTGCAGAACAAAGCAATGCGTGTCTTTGCTAG 8280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8221 CGTAATTGGAAGAGCCGCGGAGCAGAGTGCAGAACAAAGCAATGCGTGTCTTTGCTAG 8280

Qy  8281 CTGGATGAAGGTGATGGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACAGTTTGA 8340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281 CTGGATGAAGGTGATGGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACAGTTTGA 8340

Qy  8341                                                                 8400
         AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA
Db  8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400

Qy  8401 CTACTTTCTTACAAGAGATCCTGTATCCCCCTGGCAGTGCTCTGCCGCAGGGTCTGGG 8460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8401 CTACTTTCTTACAAGAGATCCTGTATCCCCCTGGCAGTGCTCTGCCGCAGGGTCTGGG 8460

Qy  8461 ATACAACCCCAGTGCTGCGTGCGTGGTATCTAATACATCACACTACCCATGTTGTGGGT 8520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8461 ATACAACCCCAGTGCTGCGTGCGTGGTATCTAATACATCACACTACCCATGTTGTGGGT 8520

Qy  8521 TAGCCGTGTGTGGCTGTCCATTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8521 TAGCCGTGTGTGGCTGTCCATTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA 8580

Qy  8581 GACTGTGACCTTTGACTGTGTGACGCGTATTGGCTTTCGGTGTGCGCTACCAACGCTGA 8640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8581 GACTGTGACCTTTGACTGTGTGACGCGTATTGGCTTTCGGTGTGCGCTACCAACGCTGA 8640

Qy  8641 CATCATTGCTGCTGTGTGTCCCAATCACTAACGACAATGACCATGCCCGCCCCCTGCGAGCCTGCGCG 8700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8641 CATCATTGCTGCTGTGTGTCCCAATCACTAACGACAATGACCATGCCCGCCCCCTGCGAGCCTGCGCG 8700

Qy  8701 GATCCTCAGAGTTCCCAATCACTAACAGACAGCCGCAGCATGACCAAGAGGCGTGGCGCCAAGAATT 8760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8701 GATCCTCAGAGTTCCCAATCACTAACAGACAGCCGCAGCATGACCAAGAGGCGTGGCGCCAAGAATT 8760

Qy  8761 AAAGAAAGCCAGGCCGGTCCTCGCCAAGAGAGGCGTGGCGCCAAGAATT 8820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8761 AAAGAAAGCCAGGCCGGTCCTCGCCAAGAGAGGCGTGGCGCCAAGAATT 8820

Qy  8821 GGCTCGCTCCTTCTCCTCTCTCTCGGCATGTCTACATCTCTAGACCTCTACAGATTGGATAAGACGAG 8880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8821 GGCTCGCTCCTTCTCCTCTCTCTCGGCATGTCTACATCTCTAGACCTCTACAGATTGGATAAGACGAG 8880

Qy  8881 CGTGGCTCGGTACACCACCTTCAATTATTGTGATGTTACTCCCCGGAGGGGGATGTGTT 8940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8881 CGTGGCTCGGTACACCACCTTCAATTATTGTGATGTTACTCCCCGGAGGGGGATGTGTT 8940

Qy  8941 TATTACACCACAGAGAAGAATTGCAGAGATTGCAGAAGTTCTGTGAAGTCATTGTCATTGTTTT 9000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8941 TGTTACACCACAGAGAAGAATTGCAGAGATTGCAGAAGTTCTGTGAAGTTTGGCTGTCATTGTTTT 9000

Qy  9001 TGCCCTAGGGCTCATTGCTGTGTTGGATTAGCCATCAGTCGAACCCCAAATTCAAAATTAA 9060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9001 TGCCCTAGGGCTCATTGCTGTGTTGGACTAGCCATCAGTCGAACCCCAAATTCAAAATTAA 9060

Qy  9061 CTAACAG---TTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC 9116
         |||||||   |||||||||||||||||||||||||||||||||||||||||||||||
Db  9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC 9120

Qy  9117 GGGCTTAACGACCCCGC 9133
         |||||||||||||||||
Db  9121 GGGCTTAACGACCCCGC 9137
```

RESULT 8
US-08-424-550B-390
; Sequence 390, Application US/08424550B
; Patent No. 6720166
; GENERAL INFORMATION:

;   APPLICANT: JOHN N. SIMONS
;   APPLICANT: TAMI J. PILOT-MATIAS
;   APPLICANT: GEORGE J. DAWSON
;   APPLICANT: GEORGE G. SCHLAUDER
;   APPLICANT: SURESH M. DESAI
;   APPLICANT: THOMAS P. LEARY
;   APPLICANT: ANTHONY SCOTT MUERHOFF
;   APPLICANT: JAMES C. ERKER

APPLICANT: SHERI L. BUIJK
APPLICANT: ISA K. MUSHAHWAR
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
NUMBER OF SEQUENCES: 716
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,550B
FILING DATE:
CLASSIFICATION: 435435
ATTORNEY/AGENT INFORMATION:
NAME: POREMBSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5527.PC.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 390:
SEQUENCE CHARACTERISTICS:
LENGTH: 9143 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-424-550B-390

Query Match 96.4%; Score 9059.8; DB 4; Length 9143;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 9096; Mismatches 37; Conservative 0; Indels 4; Gaps 1;

```
Qy    1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG   60

Qy   61 CAGGGCGTGGGGGATTTCCCCTGCGTCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CAGGGCGTGGGGGATTTCCCCTGCGTCTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  120

Qy  121 GTAGGCCGGCGGGGACTCATGACGCTCGGCCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GTAGGCCGGCGGGGACTCATGACGCTCGGCCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC  180

Qy  181 CCTGATGGGCGTTCATGGGTTCATGGGTTCGGTGGTGGTGCGCGTTAGGCAGCGCTTCCACGCCCACCA  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CCTGATGGGCGTTCATGGGTTCATGGGTTCGGTGGTGGTGCGCGTTAGGCAGCGCTTCCACGCCCACCA  240

Qy  241 CCTCCCAGATAGAGCGGCGGCGCACTGTAGGGAAGACACCGGGGACCCGGTCACTACCAAGGACG  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 CCTCCCAGATAGAGCGGCGGCGCACTGTAGGGAAGACACCGGGGACCCGGTCACTACCAAGGACG  300

Qy  301 CAGACCTCTTTTGAGTATCACGCCTCCGGAGTAGTTGGGCAAGCCCACCTATATGTGT  360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CAGACCTCTTTTGAGTATCACGCCTCCGGAGTAGTTGGGCAAGCCCACCTATATGTGT  360

Qy  361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT  420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT  420

Qy  421 CTGGGAGTCTCGTAGACCGTAGACCGTAGCCATGCACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACC  480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CTGGGAGTCTCGTAGACCGTAGACCGTAGCCATGCACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACC  480
```

```
Qy  481 TGCCCCGAGACGGCCAAGAACAAGCAGACGCAGCTTCATATCCTGTGTCCATTAAAAC  540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TGCCCCGAGACGGCCAAGAACAAGCAGACGCAGCTTCATATCCTGTGTCCATTAAAAC  540

Qy  541 ATCTGTTGAAAGGGGACAACGAGCAAAGCGCAAAGTCCAGCGCGATGCTCGGCCTCGTAA  600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 ATCTGTTGAAAGGGGACAACGAGCAAAGCGCAAAGTCCAGCGCGATGCTCGGCCTCGTAA  600

Qy  601 TTACAAAATTGCTGGTATCCATGATGGCTTGCGACACATTGGCTCAGGCTGCTTTGCCAGC  660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 TTACAAAATTGCTGGTATCCATGATGGCTTGCGACACATTGGCTCAGGCTGCTTTGCCAGC  660

Qy  661 TCATGGTTGGGACGCCAAGACCCTCGCCATAAGTCTCCCAATCTTGGAATCCTTCTGGA  720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TCATGGTTGGGACGCCAAGACCCTCGCCATAAGTCTCCCAATCTTGGAATCCTTCTGGA  720

Qy  721 TTACCCTTTGGGATTGGTGATGTTACAACTCACACACCTTCAGTAGGCCCGCTGGT  780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 TTACCCTTTGGGATTGGTGATGTTACAACTCACACACCTTCAGTAGGCCCGCTGGT  780

Qy  781 GGCAGGAGGACGGTCGTGTTCGACCAGTTCGCCAGATAGTAGCTTGCTGGGAGGATGGAGTCAA  840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GGCAGGAGGACGGTCGTGTTCGACCAGTTCGCCAGATAGTAGCTTGCTGGGAGGATGGAGTCAA  840

Qy  841 CTGGGCTACTGGTTCGGTGTCCACCTTTTTGTGGTATGTCTGCTATCTTTGGCCTG  900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CTGGGCTACTGGTTCGGTGTCCACCTTTTTGTGGTATGTCTGCTATCTTTGGCCTG  900

Qy  901 TCCCTGTAGTGGGGCGCGGGTCACTGACCAGACACAAATACCACAATCCTGACCAATTG  960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TCCCTGTAGTGGGGCGCGGGTCACTGACCAGACACAAATACCACAATCCTGACCAATTG  960

Qy  961 CTGCCAGGGTAATCAGGTTATCTATTGTTCTCCTTCCCACTTGCCTACGCAGCCTGGTTG 1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 CTGCCAGGGTAATCAGGTTATCTATTGTTCTCCTTCCCACTTGCCTACGCAGCCTGGTTG 1020

Qy 1021 TGTGATCTGCGCGGACGAGTGCTCGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 TGTGATCTGCGCGGACGAGTGCTCGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAA 1080

Qy 1081 TTGGACTGCACGGACTCCTTCTTGGCTGACCACATTGATTTTGTTATGGGCGTCTTGT 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 TTGGACTGCACGGACTCCTTCTTGGCTGACCACATTGATTTTGTTATGGGCGTCTTGT 1140

Qy 1141 GACCTGTGACGCCCTGACATTGGTGAGTTGTGTCGGTGTATTAGTCGGTGACTG 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 GACCTGTGACGCCCTGACATTGGTGAGTTGTGTCGGTGTATTAGTCGGTGACTG 1200

Qy 1201 GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGA 1260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGA 1260

Qy 1261 AGTGCCCACTGGAATAGATCCTGGGTTATCGGATGGCCGCAAGGT 1320
        |||||||||||||||||||||||||||||||||||||||||||||
Db 1261 AGTGCCCACTGGAATAGATCCTGGGTTATCGGATGGCCGCAAGGT 1320

Qy 1321 CGAGGGTGTCATCTTCTTGACCAAACTGCTTCACAAGTACCATACGCTATTGCGACTAT 1380
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 CGAGGGTGTCATCTTCTTGACCAAACTGCTTCACAAGTACCATACGCTATTGCGACTAT 1380

Qy 1381 GTTTAGCAGTGTACACTACCTGGCGTTGGCGCTCGATCTACTACTACTGGGGGCAA 1440
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 GTTTAGCAGTGTACACTACCTGGCGTTGGCGCTCGATCTACTACTACTGGGGGCAA 1440

Qy 1441 GTGGTATCAGTTGCTTCCTAGCGCTTATGCTTACATAGAAGCGACCTCTGGAAACCCCAT 1500
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 GTGGTATCAGTTGCTTCCTAGCGCTTATGCTTACATAGAAGCGACCTCTGGAAACCCTAT 1500

Qy 1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTTGCTCGCCTTTGATGATACCATGTCC 1560
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTTGCTCGCCTTTGATGATACCATGTCC 1560

Qy 1561 TTGCCACTCTTATTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC 1620
Db 1561
```

```
Db       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         TTGCCACTCTATTGGTGAGTGAGAATGTGTCAGAAGTCATTTGTTACAGTCAAAGTGGAC 1620

Qy 1621  CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621  CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG 1680

Qy 1681  TGCGAGGGGATGTATGTAAATTCAAAAATAACACATGGGTTGCTGCCGTATTCGCAA 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681  TGCGAGGGGATGTATGTAAATTCAAAAATAACACATGGGTTGCTGCCGTATTCGCAA 1740

Qy 1741  TGTGCCATCGTACTGACTATGGGCACTGATGCAGTGTGGAACGACACTCCGCAACACTTA 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741  TGTGCCATCGTACTGACTATGGGCACTGATGCAGTGTGGAACGACACTCCGCAACACTTA 1800

Qy 1801  CGAAGCATGCGGTGTAACAATGGCTAACAACCGATGGCACAACGGCTCAGCCCTGAA 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801  CGAAGCATGCGGTGTAACAATGGCTAACAACCGATGGCACAACGGCTCAGCCCTGAA 1860

Qy 1861  ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861  ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTTAAACCTCATAATTGGATGTC 1920

Qy 1921  AGGCCATTGTATTTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921  AGGCCATTGTATTTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA 1980

Qy 1981  TTCCACTCTCCTACCACCGGAGAGGTGGCTAGGTTGCCCGGTACCCCACCTGTTGTACG 2040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981  TTCCACTCTCCTACCACCGGAGAGGTGGCTAGGTTGCCCGGTACCCCACCTGTTGTACG 2040

Qy 2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTCACGTGATGATGTGAAAGACCTAGCCACAGG 2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTCACGTGATGATGTGAAAGACCTAGCCACAGG 2100

Qy 2101  ATTGATCACCAAAGACAAAGCCTGGAAAAAATTATCAGGTCTTATATTCCGGCCACGGGTGC 2160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101  ATTGATCACCAAAGACAAAGCCTGGAAAAAATTATCAGGTCTTATATTCCGGCCACGGGTGC 2160

Qy 2161  TTTGTCTCTTACGGGAGTTACCACCACCAAGGCCGTGGTTGTTTGGGGTGTGTGG 2220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161  TTTGTCTCTTACGGGAGTTACCACCACCAAGGCCGTGGTTGTTTGGGGTGTGTGG 2220

Qy 2221  CAGCAAGTAGTCTTATTTTAGCCTACTACCTCTGTGTCCCTTGTCCCTTGTTGGGGCGCTTC 2280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221  CAGCAAGTAGTCTTATTTTAGCCTACTACCTCTGTGTCCCTTGTCCCTTGTTGGGGCGCTTC 2280

Qy 2281  TGGTTACCCTTTTGCGCTCTGTGTCTCCAGTCGTATCTGTGATTTTCTTCATCGTGTTGGATGT 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2281  TGGTTACCCTTTTGCGCTCTGTGTCTCCAGTCGTATCTGTGATTTTCTTCATCGTGTTGGATGT 2340

Qy 2341  TTTGTCTAAAGCTCAAGTCACGTATCTGTGCAGCAGGTTCCTTTTCATTGGTTACGCCTGCTTCAGC 2400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2341  TTTGTCTAAAGCTCAAGTCACGTATCTGTGCAGCAGGTTCCTTTTCATTGGTTACGCCTGCTTCAGC 2400

Qy 2401  CTGCAGGCTACGTATGCTGCCCATGGGTTTGTGCCCACGGGCTGCGGGGTTGCGGACTGG 2460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401  CTGCAGGCTACGTATGCTGCCCATGGGTTTGTGCCCACGGGCTGCGGGGTTGCGGACTGG 2460.

Qy 2461  AACTTCTTTGTTGTGCAGGTTAGTTTGTGGGGCGGCGTGCTGCCCACCGATTATGACTGGGTCTGCTTGT 2520
         ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2461  AACTTCTTTGTTGTGCAGGTTAGTTTGTGGGGCGGCGTGCTGCCCACCGATTATGACTGGGTCTGCTTGT 2520

Qy 2521  AGTGGCAGGGTTAGTTTTGTGCAGCAGGTTCACCGCATAGCTCTGCTTGT 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2521  AGTGGCAGGGTTAGTTTTGTGCAGCAGGTTCACCGCATAGCTCTGCTTGT 2580

Qy 2581  AGGTCCTTGGCCTCGGTAGCGCTTTTAACCCTCTGCATTTGGTTACGCCTGCTTCAGC 2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2581  AGGTCCTTGGCCTCGGTAGCGCTTTTAACCCTCTGCATTTGGTTACGCCTGCTTCAGC 2640

Qy 2641  TTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTGAGTAGCATTAGTTGTGTCAT 2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 2641  TTTGACCGAGATAATTGGAGGGCTGACAATACCACCTGTGAGTAGCATTAGTTGTGTCAT 2700

Qy 2701  GTCTCGTTTGGCTTCTTTGCTTCACTGTTGCTCACTGTTGTAGTCTAAGTCATTAGTTAACTCCTATCT 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2701  GTCTCGTTTGGCTTCTTTGCTTCACTGTTGCTCACTGTTGTAGTCTAAGTCATTAGTTAACTCCTATCT 2760

Qy 2761  TTGGCAACGTTGGGAGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTCCT 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761  TTGGCAACGTTGGGAGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTCTCCT 2820

Qy 2821  TGTGCTGGTTTGTTCCCCGGTGCGACATATGACGGCGCTGGTGACTTTCTGTGTGTGTCA 2880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2821  TGTGCTGGTTTGTTCCCCGGTGCGACATATGACGGCGCTGGTGACTTTCTGTGTGTGTCA 2880

Qy 2881  CGTAGCTCTTCTATGTTTAACATCCAGTGCGCAGCAGGCACAACGGCTCAGCCCTGAA 2940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2881  CGTAGCTCTTCTATGTTTAACATCCAGTGCGCAGCAGGCACAACGGCTCAGCCCTGAA 2940

Qy 2941  TAGGGCCCATAGAAATGTTGGTGCGTCTCGGAAAGTGTCTATTCTTGGTATTCTCATTATGT 3000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2941  TAGGGCCCATAGAAATGTTGGTGCGTCTCGGAAAGTGTCTATTCTTGGTATTCTCATTATGT 3000

Qy 3001  TCTTAAGTTTTCCTCTTAGTGTTGTGTTGGTTGCCTCTAATGATTTTGCTGAAACTACCCTGTGGGGACAC 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3001  TCTTAAGTTTTCCTCTTAGTGTTGTGTTGGTTGCCTCTAATGATTTTGCTGAAACTACCCTGTGGGGACAC 3060

Qy 3061  TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTTTTCCC 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 3061  TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTTTTCCC 3120

Qy 3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACGCTTGGCGTGTGGGGACAC 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACGCTTGGCGTGTGGGGACAC 3180

Qy 3181  GGTTGATGGTTTGCCCGTTGTTGCCCGTTGTGTTGCTCGACACCTTTTACGCTGTCTTCGGCGGTGC 3240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db 3181  GGTTGATGGTTTGCCCGTTGTTGCCCGTTGTGTTGCTCGACACCTTTTACGCTGTCTTTAGCTAT 3240

Qy 3241  GCCGCCAGATGGGTGGGCCATTACCGCGTTTTACGCTGTCTCTGAACGTGG 3300
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3241  GCCGCCAGATGGGTGGGCCATTACCGCGTTTTACGCTGTCTCTGAACGTGG 3300

Qy 3301  CACGGCTGTCAGCGATGGCAGTGCAGTTGTATAGACCCCGAACTTGGAACTGGAAC 3360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3301  CACGGCTGTCAGCGATGGCAGTGCAGTTGTATAGACCCCGAACTTGGAACTGGAAC 3360

Qy 3361  TATCTTCAGATTAGGATCTCTCGGCCACTAGCTACTACACATGGTGGCTCATCCCAGGCGTGTT 3420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3361  TATCTTCAGATTAGGATCTCTCGGCCACTAGCTACTACACATGGTGGCTCATCCCAGGCGTGTT 3420

Qy 3421  GTATACTGCTCCACCATGCGCAGCAAGGGGCGCCGGTTAATGACCAACCACCATGTGGACAACGTGTT 3480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3421  GTATACTGCTCCACCATGCGCAGCAAGGGGCGCCGGTTAATGACCAACCACCATGTGGACAACGTGTT 3480

Qy 3481  CCCAATAACCGGTTGACGCGGCTAATGCGTGACGCGGCTAATGACCAACCACCATGTGGAGACTGG 3540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3481  CCCAATAACCGGTTGACGCGGCTAATGCGTGACGCGGCTAATGACCAACCACCATGTGGAGACTGG 3540

Qy 3541  GTCCCTACTCGGTGCTCTTGGGGGAGACCAAGGGTATCGGTGTGGGGTC 3600
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 3541  GTCCCTACTCGGTGCTCTTGGGGGAGACCAAGGGTATCGGTGTGGGGTC 3600

Qy 3601  ATTGGTTGAGGTCAACAAATCCGATGAGCTCTATTGGTGTGTGCGGGGCCCTTCCCAT 3660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3601  ATTGGTTGAGGTCAACAAATCCGATGAGCTCTATTGGTGTGTGCGGGGCCCTTCCCAT 3660

Qy 3661  GGCTGTTGCCAAGGGTTCTTCAGGTGCCGATTCTGTGCTCCTCCCGGGCATGTTATTGG 3720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3661  GGCTGTTGCCAAGGGTTCTTCAGGTGCCGATTCTGTGCTCCTCCCGGGCATGTTATTGG 3720

Qy 3721  GATGTTCACCGCTGCTAGAAATTCTGCGGCGGTTCAGTCAGTCAGAATTAGGGTTAGGCCGTT 3780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3721  GATGTTCACCGCTGCTAGAAATTCTGCGGCGGTTCAGTCAGTCAGCCAGATTAGGGTTAGGCCGTT 3780
```

```
Qy  3781 GGTGTGTGCTGATACCATCCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC 3840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781 GGTGTGTGCTGATACCATCCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC 3840

Qy  3841 TGTGCCTAACGAGTATTCAGTGACGCATGTCAAATTTAATTGCCCCCACTACTGCAAGTCAAC 3900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 TGTGCCTAACGAGTATTCAGTGACGCATGTCAAATTTAATTGCCCCCACTACTGCAAGTCAAC 3900

Qy  3901 CAAATTACCACTTCTTACATGCAGGAGAAGTATGAGGTCTTGGTCCTTAAATCCCAGTGT 3960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 CAAATTACCACTTCTTACATGCAGGAGAAGTATGAGGTCTTGGTCCTTAAATCCCAGTGT 3960

Qy  3961 GGCTACAACAGCATCAATGCCAAAGTACATGCACGCGACGTACGCGACGGTGAATCCAAATTG 4020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 GGCTACAACAGCATCAATGCCAAAGTACATGCACGCGACGTACGCGACGGTGAATCCAAATTG 4020

Qy  4021 CTATTTTAATGCAAAGTACCAACACAGGGGCTTCACTTACGTACGCACAGCACACATATGGCAT 4080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 CTATTTTAATGCAAAGTACCAACACAGGGGCTTCACTTACGTACGCACAGCACACATATGGCAT 4080

Qy  4081 GTACCTGACCGGAGCATGTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGC 4140
         ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db  4081 GTACCTGACCGGAGCATGTCCCGGAACTATGATGTCATCATTTGTGACGAATGCCATGC 4140

Qy  4141 TACCGATGCAACCACCGTGTTGGGCATTGGAAAGGTCCTAACCGAAGCTCCATCCAAAAA 4200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 TACCGATGCAACCACCGTGTTGGGCATTGGAAAGGTTCCTAACCGAAGCTCCATCCAAAAA 4200

Qy  4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCCTGGAGTAATCCCTACCTACCACATGC 4260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCCTGGAGTAATCCCTACCTACCACATGC 4260

Qy  4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAGAT 4320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261 CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAGAT 4320

Qy  4321 TAAGGAGAAATCTGAAGAAGGGAGGAGACACCTTATCTTTGAGGCTACCAAAAAACACTG 4380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321 TAAGGAGAAATCTGAAGAAGGGAGGAGACACCTTATCTTTGAGGCTACCAAAAAACACTG 4380

Qy  4381 TGATGAGCTTGCTGCTAACGAGTTAGCTCGAAAGGAATAACAGCTGTCTCTTACTATAGGGG 4440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381 TGATGAGCTTGCTGCTAACGAGTTAGCTCGAAAGGAATAACAGCTGTCTCTTACTATAGGGG 4440

Qy  4441 ATGTGACATCTCAAAAATCCCTGAGGGGCGACTGTGTAGTTCCTGAACATTGTTGAAGCCTT 4500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441 ATGTGACATCTCAAAAATCCCTGAGGGGCGACTGTGTAGTTCCTGAATGTCAACATTGTTGAAGCCTT 4500

Qy  4501 TACAGGGTACACTGGTACTTGACCTTGATTCCGTGTATGACTGCAGCCTCATGTGATAGAAGGCAC 4560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501 TACAGGGTACACTGGTACTTGACCTTGATTCCGTGTATGACTGCAGCCTCATGTGATAGAAGGCAC 4560

Qy  4561 ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTGTTCGTGTGTCGGGGTTCAGC 4620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561 ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTGTTCGTGTGTCGGGGTTCAGC 4620

Qy  4621 AATAGTTAAAGGCCAGCGGCCGACACAGGCCGCATATACTACTA 4680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621 AATAGTTAAAGGCCAGCGGCCGACACAGGCCGCATATACTACTA 4680

Qy  4681 TGTGGTTTAGTACTCTCCGGTGTACCCCCTTCGGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681 TGCTGGTTTAGTACTCTCCGGTGTACCCCCTTCGGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740

Qy  4741 CGACGCAGCCAAGGCCAAGGCCACCCAACCTGGGTTACCTGCAGCCAAATTTGGACGAGTGGCTGATCT 4800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741 CGACGCAGCCAAGGCCAAGGCCACCCAACCTGGGTTACCTGCAGCCAAATTTGGACGAGTGGCTGATCT 4800

Qy  4801 CTATCGGCACCCAACCTGGGTTACCTGCAGCCAAATTTGGACGAGTGGCTGATCT 4860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801 CTATCGGCACCCAACCTGGGTTACCTGCAGCCAAATTTGGACGAGTGGCTGATCT 4860
```

```
Qy  4861 CTTTTCTATGGTCAACCCCGAACTTCATTGTCAATACTGCAAATAAAGAACTGCTGACAA 4920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861 CTTTTCTATGGTCAACCCCGAACTTCATTGTCAATACTGCAAATAAAGAACTGCTGACAA 4920

Qy  4921 TTATGTTTGTTGTTGACTGAGCGCAGGGAGGAGCCCGGCTGCATCAGTATGGCTCATGCTGCTCC 4980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921 TTATGTTTGTTGTTGACTGAGCGCAGGGAGGAGCCCGGCTGCATCAGTATGGCTCATGCTGCTCC 4980

Qy  4981 CAATGACGCCACCACGGTGGCAGGGAGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG 5040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981 CAATGACGCCACCACGGTGGCAGGGAGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCTGTG 5040

Qy  5041 GCGCTTGGACGGGCCTGACGCCTGCTCCTGCCCCAGAGCCCAGCCAGGTGACCAGATACCA 5100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041 GCGCTTGGACGGGCCTGACGCCTGCTCCTGCCCCAGAGCCCAGCCAGGTGACCAGATACCA 5100

Qy  5101 AATGTGCTTCACTGAAGTCAATACTTCTGGGACAGCCCGGCACTCGCTGTGTGGCGGTTGGAGT 5160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101 AATGTGCTTCACTGAAGTCAATACTTCTGGGACAGCCCGGCACTCGCTGTGTGGCGGTTGGAGT 5160

Qy  5161 GGCTATGGCTTATCTAGCAGCCATTGACACACTTTGGCGCGGCCACTTGTGTGCGGCGTTGCTGGTC 5220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161 GGCTATGGCTTATCTAGCAGCCATTGACACACTTTGGCGCGGCCACTTGTGTGCGGCGTTGCTGGTC 5220

Qy  5221 TATTACATCAGTCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAAGAAATCGT 5280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221 TATTACATCAGTCCTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAAGAAGAAATCGT 5280

Qy  5281 GGAGGAGTGTGATCATTCATTCCCTTGGAGGCCATGGTGTTGCTGCAATTGACAAGCTGAA 5340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281 GGAGGAGTGTGATCATTCATTCCCTTGGAGGCCATGGTGTTGCTGCAATCGATAAGCTGAA 5340

Qy  5341 GAGTACAATCACCACACAGTCCTTCACATTGGAAACCGCCTTGAAAAACTTAACAC 5400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341 GAGTACAATAACCACCACAGTCCTTCACATTGGAAACCGCCTTGAAAAACTTAACAC 5400

Qy  5401 CTTTCTTGGGCCTCATGCAGCTACAAATCCTTGCTACATCATAGATATTGCTGTGGTTTAGT 5460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401 CTTTCTTGGGCCTCATGCAGCTACAAATCCTTGCTACATCATAGATATTGCTGTGGCTTAGT 5460

Qy  5461 CACTTTACCTGACAATCCCTTTGCATCGGTGTTTGCTTTCATTGCGGGTTATTACTAC 5520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461 CACTTTACCTGACAATCCCTTTGCATCGGTGTTTGCTTTCATTGCGGGTTATTACTAC 5520

Qy  5521 CCCACTACCTCACAAGATCAAAATGTTCCTGCTACATTATTGGAGGCGCAATTGCGTCCAA 5580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521 CCCACTACCTCACAAGATCAAAATGTTCCTGCTACATTATTGGAGGCGCAATTGCGTCCAA 5580

Qy  5581 GCTTACAGACGCTAGAGGCGCACTGGCGTTCATGATGCGCTCGGCGGGGCTCGCGGGAACAGCTCT 5640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581 GCTTACAGACGCTAGAGGCGCACTGGCGTTCATGATGCGCTCGGCGGGGCTCGCGGGAACAGCTCT 5640

Qy  5641 TGGTACATGGACATCGGTGGGTTTTGTCTTTGACATGTGACATGCTAGGCGGCTATGCTCGCCGCCTC 5700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641 TGGTACATGGACATCGGTGGGTTTTGTCTTTGACATGTGACATGCTAGGCGGCTATGCTCGCCGCCTC 5700

Qy  5701 ATCCACTGCTTGCTTGACATTTAAATGCTTAAATGCTTGACATTTAAATGCTTGATCAGCT 5760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701 ATCCACTGCTTGCTTGACATTTAAATGCTTAAATGCTTGACATTTAAATGCTTGATCAGCT 5760

Qy  5761 TGCTGGTTTAGTCTACTCCGGTCAATCGGCCCAGGAGTTGTGGGCGTCTTGTCAGC 5820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761 TGCTGGTTTAGTCTACTCCGGTTCAATCGGCCCAGGAGTTGTGGGCGTCTTGTCAGC 5820

Qy  5821 TTGTGCAATGTTTGCTTTTGTTTTGTGACAACACTGCTACTTCTTACACTTCTTAC 5880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821 TTGTGCAATGTTTGCTTTTGTTTTGTGACAACACTGCTACTTCTTACACTTCTTAC 5880

Qy  5881 TAATGCTTGCTAGGAGCAACACTGTATGTGTAATGAGTACTTTATTGCCACTCGGTACATCCG 5940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881 TAATGCTTGCTAGGAGCAACACTGTATGTGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940

Qy  5941 CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCTCGGAGTGTCATATCAGCTTGCAT 6000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941 CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCTCGGAGTGTCATATCAGCTTGCAT 6000
```

Db

Qy 7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAACAGGCGGAGGCCCTGATGA 7080

Db

Qy 7081 TTTACCCAGTACCCTCCCAAAAAGAGGAGGTCTCTGAATGGTCAGACGAAAGTGGTCGAC 7140
Db 7081 TTTACCCAGTACCCTCCCAAAAAGAGGAGGTCTCTGAATGGTCAGACGAAGTGGTCAAC 7140

Qy 7141 GGCTACAACCGCTTCCAGCTACGTTACTGGCCCCCCGTACCCTAAGATACGGGGAAAGGA 7200
Db 7141 GACTACAACCGCTTCCAGCTACGTTACTGGCCCCCCGTACCCTAAGATACGGGGAAAGGA 7200

Qy 7201 TTCCACTCAGTCAGCCCCGCCAAACCGCCTACAAAAGAAGTGGGAAAAGAGTGAGTT 7260
Db 7201 TTCCACTCAGTCAGCCCCGCCAAACCGCCTACAAAAGAAGTGGGAAAGAGTGAGTT 7260

Qy 7261 TTCGTGAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320
Db 7261 TTCGTGAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320

Qy 7321 TCTGTCTGCAACTCGGGCCATCACTACGTGGTTCCTCAAACAAAGATCATTGGTGTATGT 7380
Db 7321 TCTGTCTGCAACTCGGGCCATCACTACGTGGTTCCTCAAACAAAGATCATTGGTGTATGT 7380

Qy 7381 GACTGAGCCGCCGGGATGCGGACTGGAGCTTAGAAAACAAAAGTCACTATTAATAGACAACCTCT 7440
Db 7381 GACTGAGCCGCCGGGATGCGGACTGGAGCTTAGAAAACAAAAGTCACTATTAATAGACAACCTCT 7440

Qy 7441 GTTCCCCCATCATCACACAAGTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500
Db 7441 GTTCCCCCATCATCACACAAGTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500

Qy 7501 CGGTGTCATGTGGGACTATGATGAAGTAGCAGCTCACGCGCCCTCTAAGTCTGCTAAGTC 7560
Db 7501 CGGTGTCATGTGGGACTATGATGAAGTAGCAGCTCACGCGCCCTCTAAGTCTGCTAAGTC 7560

Qy 7561 CACACTACTGGCCTTCGGGGCACTGATGTTCGTTCGTTCTGGAGCAGCCCGCAAGGCTGTTCT 7620
Db 7561 CACACTACTGGCCTTCGGGGCACTGATGTTCGTTCGTTCTGGAGCAGCCCGCAAGGCTGTTCT 7620

Qy 7621 GGACTTGCAGAGATGTGTGTCGAGGCAGGTGAGATACCGGTAAACTGTGAT 7680
Db 7621 GGACTTGCAGAGATGTGTGTCGAGGCAGGTGAGATACCGGTAAACTGTGAT 7680

Qy 7681 AGTTCCAAAGAGGAGGTCTTCGTGGTGTGTCGAAGACCCCCAGAAACCAACAAGAAACCCCCAAG 7740
Db 7681 AGTTCCAAAGAGGAGGTCTTCGTGGTGTGTCGAAGACCCCCAGAAACCAACAAGAAACCCCCAAG 7740

Qy 7741 GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGT 7800
Db 7741 GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGT 7800

Qy 7801 TGCTCCTGACGTAGTTAAAGCGTCATGGGAGATGCTGGTTTGTCGACCCACGTAC 7860
Db 7801 TGCTCCTGACGTAGTTAAAGCGTCATGGGAGATGCTGGTTTGTCGACCCACGTAC 7860

Qy 7861 CCGTGTCAAGCGTCTGTTGTCCGATGTCGTCACCCGGATGCAGTCGAGCCACATGCGATAC 7920
Db 7861 CCGTGTCAAGCGTCTGTTGTCCGATGTCGTCACCCGGATGCAGTCGAGCCACATGCGATAC 7920

Qy 7921 AGTGTGTTTGACAGTACCATCACACCCGAGATATCATGGTGAGACAGACATCTACTC 7980
Db 7921 AGTGTGTTTGACAGTACCATCACACCCGAGATATCATGGTGAGACAGACATCTACTC 7980

Qy 7981 AGCAGCTAAACTCAGTGACCAACACCAGCTGGCAATTCACACCATTGCGAGGCAGTTATA 8040
Db 7981 AGCAGCTAAACTCAGTGACCAACACCAGCTGGCAATTCACACCATTGCGAGGCAGTTATA 8040

Qy 8041 CGCTGAGGACCGATGCTTATGAGTGGCCGAGAGATCGGATATCGTAGGTGTAGGTC 8100
Db 8041 CGCTGAGGACCGATGCTTATGAGTGGCCGAGAGATCGGATATCGTAGGTGTAGGTC 8100

Qy 8101 TTCCGGCGTCTATACTACCTCAAGTTCCAACAGTTTGACCTGCTGGCTGCGTGAAGTAAATGC 8160
Db 8101 TTCCGGCGTCTATACTACCTCAAGTTCCAACAGTTTGACCTGCTGGCTGCGTGAAGTAAATGC 8160

Db

Qy 5941 CAGGAAGATACTGGGGCATTCTGGAGGCATCTACCCCCTGGAGTGTGTCATATCAGCTTGCAT 6000

Db

Qy 6001 CCGTTGGCTCCACACCCCGACGGAGGAGATGATTGCGGCCTCATTGCTTGGGGTCTCAGAGAT 6060
Db 6001 CCGTTGGCTCCACACCCCGACGGAGGAGATGATTGCGGCCTCATTGCTTGGGGTCTCAGAGAT 6060

Qy 6061 TTGGCAGTATGTGTGCAATTCTTCTTGTGATTGCTTTAAATGTCCTTAAAGCTGGAGTTCA 6120
Db 6061 TTGGCAGTATGTGTGCAATTCTTCTTGTGATTGCTTTAAATGTCCTTAAAGCTGGAGTTCA 6120

Qy 6121 GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC 6180
Db 6121 GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCC 6180

Qy 6181 CTGGAATTGGATCAGGTATGCTCCAAGCACGCGTGTCCATGCGGTCCATTGCTGAACTCATCTTTC 6240
Db 6181 CTGGAATTGGATCAGGTATGCTCCAAGCACGCGTGTCCATGCGGTCCATTGCTGAACTCATCTTTC 6240

Qy 6241 TGTTGAGAAATGTTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG 6300
Db 6241 TGTTGAGAAATGTTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG 6300

Qy 6301 AGGGGCTGTGTTCCAGTCAACGCGTAGGCTGTGTGGGTCGGCTAGACCGGGTACAAGGTGTGATTG 6360
Db 6301 AGGGGCTGTGTTCCAGTCAACGCGTAGGCTGTGTGGGTCGGCTAGACCGGGTACAACTGATTG 6360

Qy 6361 GACTAGTTCTTGTCGTCAATTATGGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA 6420
Db 6361 GACTAGTTCTTGTCGTCAATTATGGGCGTTAGGGACTACTGTAAATATGAGAAATTGGGAGA 6420

Qy 6421 TCACATTTTTGTTGTTACAGCAGTATCCTCTCCAAAATGTCTGTTTCACCCAGTGCCCCCAAC 6480
Db 6421 TCACATTTTTGTTGTTACAGCAGTATCCTCTCCAAAATGTCGTTTCACCCAGTGCCCCCAAC 6480

Qy 6481 CTTGAGAGCTGCAGTGGCCGTGAGCGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540
Db 6481 CTTGAGAGCTGCAGTGGCCGTGAGCGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540

Qy 6541 AACTCCTTGGACGACATCTGCTTGCTGTGTTACGGTCCTGACGGTAAAACTGTTAA 6600
Db 6541 AACTCCTTGGACGACATCTGCTTGCTGTGTTACGGTCCGGACGGTAAGGGTAAAACTGTTAA 6600

Qy 6601 GCTTCCCTTCCGCGTTGACGGTCACACACTGGTGTCGGCGCATGCAACTTAATTTGCGGTGA 6660
Db 6601 GCTTCCCTTCCGCGTTGACGGTCACACACTGGTGTCGGCGCATGCAACTTAATTTGCGGTGA 6660

Qy 6661 TGCACTTGAGACAAATGACTGGTTCGCTAATTCCACAACACACTCCTAGTGATGAGCCGCAGT 6720
Db 6661 TGCACTTGAGACAAATGACTGGTTCGCTAATTCCATAAACAACACTCCTAGTGATGAGCCGCAGT 6720

Qy 6721 GTCCGGCTCTTGTTTTCAAAACGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC 6780
Db 6721 GTCCGGCTCTTGTTTTCAAAACGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC 6780

Qy 6781 CCCTCCTTCCACCACCTGTTCTTACAGTGGCCATGCCAGCCCCCTTGTTGGGAGGTAGTGGTTAAGAAA 6840
Db 6781 CCCTTCTTCCACCACCTGTTCTTACAGTGGCCATGCCAGCCCCCTTGTTGGGAGGTAGTGGTTAAGAAA 6840

Qy 6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTGCCTTCGCCTCCCCCCCGAGATCCCC 6900
Db 6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTGCCTTCGCCTCCCCCCCGAGATCCGTCCC 6900

Qy 6901 AGGAGGTGTCATGTGGTCCTGAAAGCCTGCAACGAAGTGACCGTTAGAAGGTCCTTCAAACCT 6960
Db 6901 AGGAGGTGTCATGTGGTCCTGAAAGCCTGCAACGAAGTGACCGTTAGAAGGTCCTTCAAACCT 6960

Qy 6961 CCCTTCTTCACCACCTGTTCTTACAGTTGGGCCATGCGCCGGGTGA 7020
Db 6961 CCCTTCTTCACCACCTGTTCTTACAGTTGGGCCATGCGCCGGGTGA 7020

Qy 7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAGGCGGAGGCCCTGATGA 7080
Db 7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAACAGGCGGAGGCCCTGATGA 7080

```
Patent No. 6720166
GENERAL INFORMATION:
    APPLICANT: JOHN N. SIMONS
    APPLICANT: TAMI J. PILOT-MATIAS
    APPLICANT: GEORGE J. DAWSON
    APPLICANT: GEORGE G. SCHLAUDER
    APPLICANT: SURESH M. DESAI
    APPLICANT: THOMAS P. LEARY
    APPLICANT: ANTHONY SCOTT MUERHOFF
    APPLICANT: JAMES C. ERKER
    APPLICANT: SHERI L. BUIJK
    APPLICANT: ISA K. MUSHAHWAR
    TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
    TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
    NUMBER OF SEQUENCES: 716
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: ABBOTT LABORATORIES    D377/AP6D
        STREET: 100 ABBOTT PARK ROAD
        CITY: ABBOTT PARK
        STATE: IL
        COUNTRY: USA
        ZIP: 60064-3500
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/424,550B
        FILING DATE:
        CLASSIFICATION: 435435
    ATTORNEY/AGENT INFORMATION:
        NAME: POREMBSKI, PRISCILLA E.
        REGISTRATION NUMBER: 33,207
        REFERENCE/DOCKET NUMBER: 5527.PC.01
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: 708-937-6365
        TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 393:
    SEQUENCE CHARACTERISTICS:
        LENGTH: 9143 base pairs
        TYPE: nucleic acid
        STRANDEDNESS: double
        TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
    FEATURE:
        NAME/KEY: 5'UTR
        LOCATION: 1..445
    FEATURE:
        NAME/KEY: CDS
        LOCATION: 446..9037
    FEATURE:
        NAME/KEY: 3'UTR
        LOCATION: 9038..9143
US-08-424-550B-393
```

```
Query Match    96.4%;    Score 9059.8;    DB 4;    Length 9143;
Best Local Similarity 99.6%;    Pred. No. 0;
Matches 9096;    Conservative 0;    Mismatches 37;    Indels 4;    Gaps 1;

Qy      1 ACCACAAACACTCCAGTTTGTTACACTCCGTAGGAATGCTCCTGGAGCACCCCCCTAG  60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACCACAAACACTCCAGTTTGTTACACTCCGTAGGAATGCTCCTGGAGCACCCCCCTAG  60

Qy     61 CAGGGCGTGGGGGATTTCCCCTGCCCGTCTGCGAAAGGGTGGAGCCAACCACCTTAGTAT  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAGGGCGTGGGGGATTTCCCCTGCCCGTCTGCGAAAGGGTGGAGCCAACCACCTTAGTAT  120

Qy    121 GTAGGCGGCGGGACTCATGACGCTCGCGTGATGACAAGCGCGCCAAGCTTGACTTGGATGGC  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTAGGCGGCGGGACTCATGACGCTCGCGTGATGACAAGCGCGCCAAGCTTGACTTGGATGGC  180

Qy    181 CCTGATGGCGGTTCATGGGTTCGGTGGTGGCTTTAGGCAGCGCTCCACGCCCACCA  240
```

(Left column — sequence alignment, Qy/Db pairs)

```
Qy  8161 TGCAGCCGAACAGGCTGGCATGAAGAACCCTCGCTTCCTTATTTGCGGCGATGATTGCAC  8220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161 TGCAGCCGAACAGGCTGGCATGAAGAACCCTCGCTTCCTTATTTGCGGCGATGATTGCAC  8220

Qy  8221 CGTAATTTGGAAGAGCGCCGGAGCGCAGATGCGACAAACAAGCAATGCGTGTCTTGCTAG  8280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8221 CGTAATTTGGAAGAGCGCCGGAGCGCAGATGCGACAAACAAGCAATGCGTGTCTTGCTAG  8280

Qy  8281 CTGGATGAAGGTGATGGGTGCACCACAGATTGTGTGCCTCAACCCAAATACAGTTTGGA  8340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281 CTGGATGAAGGTGATGGGTGCACCACAGATTGTGTGCCTCAACCCAAATACAGTTTGGA  8340

Qy  8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA  8400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA  8400

Qy  8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTGGCCAGGTGCTCTGCCGAGGGTCTCGGG  8460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTGGCCAGGTGCTCTGCCGAGGGTCTCGGG  8460

Qy  8461 ATACAACCCAGTGCGCTGGGATTGGGTATCTAATACATCACTACCCCATGTTTGTGTGGT  8520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8461 ATACAACCCAGTGCGCTGGGATTGGGTATCTAATACATCACTACCCCATGTTTGTGTGGT  8520

Qy  8521 TAGCCGTGTGTGTGGCTGTCCATTTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA  8580
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8521 TAGCCGTGTGTGTGGCTGTCCATTTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA  8580

Qy  8581 GACTGTGACCTTGACTGGTATGGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG  8640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8581 GACTGTGACCTTGACTGGTATGGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG  8640

Qy  8641 CATCATTGCTGCGTGTGCACGGTATTGGAGGCTTTCTCGGTGGTGCGCTACACCAACGCTGA  8700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8641 CATCATTGCTGCGTGTGCACGGTATTGGAGGCTTTCTCGGTGGTGCGCTACACCAACGCTGA  8700

Qy  8701 GATCCTCAGAGTTTCCCAATCACTAACAGACAACATGACCATGCCCCCCTGCGAGCCTGGCG  8760
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8701 GATCCTCAGAGTTTCCCAATCACTAACAGACAACATGACCATGCCCCCCTGCGAGCCTGGCG  8760

Qy  8761 AAAGAAAGCCAGGCGTCCTGCCCAGCGGCGTGGAGCGAGAGGCGTGCACACGCAAAATT  8820
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8761 AAAGAAAGCCAGGCGTCCTGCCCAGCGGCGTGGAGCGAGAGGCGTGCACACGCAAAATT  8820

Qy  8821 GGCTTCGCTTCCTTCTCTCGGCATGCTACATCTAGACCTTCTACCAGATTGGATAAGACGAG  8880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8821 GGCTTCGCTTCCTTCTCTCGGCATGCTACATCTAGACCTTCTACCAGATTGGATAAGACGAG  8880

Qy  8881 CGTGGCTTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGGAGGGGGATGTGTT  8940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8881 CGTGGCTTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGGAGGGGGATGTGTT  8940

Qy  8941 TATTACACCACAGAGAAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTTT  9000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8941 TGTTACACCACAGAGAAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTTT  9000

Qy  9001 TGCCCTAGGGCTCATTGCTTGTGTTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAA  9060
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9001 TGCCCTAGGGCTCATTGCTTGTGTTGGACTAGCCATCAGCTGAACCCCAAATTCAAAATTAA  9060

Qy  9061 CTAACAG---TTTTTTTTTTTTTTTTTTTTTAGGCGGCAGCGCAACAGGGGAGACCCC  9116
         |||||||   |||||||||||||||||||||||||||||||||||||||||||||||||
Db  9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTAGGCGGCAGCGCAACAGGGGAGACCCC  9120

Qy  9117 GGGCTTAACGACCCCGC  9133
         |||||||||||||||||
Db  9121 GGGCTTAACGACCCCGC  9137
```

```
RESULT 9
US-08-424-550B-393
; Sequence 393, Application US/08424550B
```

```
Db  1261  AGTGCCACTGGAATAGATCTCGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT  1320
Qy  1321  CGAGGCTGTCATCCTTCTTGACCAAACTGGCTTCACAAGTACCATTGCGACTAT  1380
Db  1321  CGAGGCTGTCATCCTTCTTGACCAAACTGGCTTCACAAGTACCATTGCGACTAT  1380
Qy  1381  GTTTAGCAGTGTACACTACCTGGCGCTTGGCGCTCTGATCTACTACTATATGTGT  1440
Db  1381  GTTTAGCAGTGTACACTACCTGGCGCTTGGCGCTCTGATCTACTACTATATGTGT  1440
Qy  1441  GTGGTATCAGTTGCTCCTAGCGGTTACATGCTTTACATAGAAGGGACCTCTGGAAACCCAT  1500
Db  1441  GTGGTATCAGTTGCTCCTAGCGGTTACATGCTTTACATAGAAGGGACCTCTGGAAACCCTAT  1500
Qy  1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGTCGCCTTGATGATACCATGTCC  1560
Db  1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGTCGCCTTTGATGATACCATGTCC  1560
Qy  1561  TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTGCAAAGTGGAC  1620
Db  1561  TTTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTGCAAAGTGGAC  1620
Qy  1621  CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCTGG  1680
Db  1621  CAGGCCTGTCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG  1680
Qy  1681  TGGCGAGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA  1740
Db  1681  TGGCGAGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCAA  1740
Qy  1741  TGTGCCATCGTACTGCACTATGGGACTGATGCGAGTGTGGAACGACACTCGCAACACTTA  1800
Db  1741  TGTGCCATCGTACTGCACTATGGGACTGATGCGAGTGTGGAACGACACTCGCAACACTTA  1800
Qy  1801  CGAAGCATGCGGTGTAACAACATGGCTAACAACCGCATGGCACAACGGCTCAGCCTGAA  1860
Db  1801  CGAAGCATGCGGTGTAACAACATGGCTAACAACCGCATGGCACAACGGCTCAGCCTGAA  1860
Qy  1861  AGTGGCTATATTACAATACCCTGGGTCTAAAGAAAATGTTAAACCTCATAATTGGATGTC  1920
Db  1861  ATTGGCTATATTACAATACCCTGGGTCTAAAGAAAATGTTAAACCTCATAATTGGATGTC  1920
Qy  1921  AGGCCATTGTATTTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA  1980
Db  1921  AGGCCATTGTATTTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA  1980
Qy  1981  TTCCACTTCTCCTACCGGAGAGTGGCTAGGTTGCCGGTACCCCACCTGTGGTACG  2040
Db  1981  TTCCACTTCTCCTACCGGAGAGTGGCTAGGTTGCCGGTACCCCACCTGTGGTACG  2040
Qy  2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAGG  2100
Db  2041  ATTGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAGG  2100
Qy  2101  ATTGATCACCAAAGACAAAGCCTGAAAAATTATCAGGTCTTATATTCCGCCACGGGGTGC  2160
Db  2101  ATTGATCACCAAAGACAAAGCCTGAAAAATTATCAGGTCTTATATTCCGCCACGGGGTGC  2160
Qy  2161  TTTGTCTTCTTACGGAGTTACCACCAAGGCCGTGGTGCTAATTCTCGTTGGGGTTGTGTGG  2220
Db  2161  TTTGTCTTCTTACGGAGTTACCACCAAGGCCGTGGTGCTAATTCTCGTTGGGGTTGTGTGG  2220
Qy  2221  CAGCAAGTGATCTTATTTTAGCCTCCGTGTCGTTTTGATTTTCTTCATCGTGTCATCTCCG  2280
Db  2221  CAGCAAGTGATCTTATTTTAGCCTCCGTGTCGTTTTGATTTTCTTCATCGTGTCATCTCCG  2280
Qy  2281  TGGTTACCCTTTGCGTCCTGTCTCCATCCCAAGCTGTATCCTCGTATCTTGTTGTTGTGG  2340
Db  2281  TGGTTACCCTTTGCGTCCTGTCTCCATCCCAAGCTGTATCCTCGTATCTTGTTGTTGTGG  2340
Qy  2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTTTGTTCTTTCATCGTGTCATCTCCG  2400
Db  2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTTTGTTCTTTCATCGTGTCATCTCCG  2400
```

```
Db  181   CCTGATGGGCGGTTCATGGGTTCGGTGGTGGTGGCGGCCTTAGGCAGCCTCCACGCCCACCA  240
Qy  241   CCTCCCAGATAGAGCGGCGGCCACTGTAGGGAGGAAGACCGGGACCGGTCACTACCAAGGACG  300
Db  241   CCTCCCAGATAGAGCGGCGGCCACTGTAGGGAGGAAGACCGGGACCGGTCACTACCAAGGACG  300
Qy  301   CAGACCTCCTTTTGAGTATCACGCCTCCGGAAGTAGTTGGGCAAGCCACCTATATGTGT  360
Db  301   CAGACCTCCTTTTGAGTATCACGCCTCCGGAAGTAGTTGGGCAAGCCACCTATATGTGT  360
Qy  361   TGGGATGGTTGGGGTTAGCCATCCATCATACCGTACTGCCTGATAGGGTCCTTGCTGAGGGGAT  420
Db  361   TGGGATGGTTGGGGTTAGCCATCCATCATACCGTACTGCCTGATAGGGTCCTTGCTGAGGGGAT  420
Qy  421   CTTGGGAGTCTCGTAGACCGTAGCACATGGCGTGTATTTCTACTCAAACAAGTCCTGTACC  480
Db  421   CTTGGGAGTCTCGTAGACCGTAGCACATGGCGTGTATTTCTACTCAAACAAGTCCTGTACC  480
Qy  481   TGCGCCCAGAACGCGCAAGAACAAGCAGCAGGCTTCATATCCTGTGTCCATTAAAAC  540
Db  481   TGCGCCCAGAACGCGCAAGAACAAGCAGCAGGCTTCATATCCTGTGTCCATTAAAAC  540
Qy  541   ATCTGTTGAAAGGGGACGGCAAAGCGCAAAGTCAGCGCGATGCTCGGCCCTTCGTAA  600
Db  541   ATCTGTTGAAAGGGGACGGCAAAGCGCAAAGTCAGCGCGATGCTCGGCCCTTCGTAA  600
Qy  601   TTACAAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC  660
Db  601   TTACAAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC  660
Qy  661   TCATGGTTGGGGACGGCCAAGACCCTGCCAATAAGTCTCGCAATCTTCGGAATCTTCTGA  720
Db  661   TCATGGTTGGGGACGGCCAAGACCCTGCCAATAAGTCTCGCAATCTTCGGAATCTTCTGA  720
Qy  721   TTTACCCCTTTGGGGGTGATTGGTGATGTTACAACTCACACACTCTAGTAGGCCGCTGGT  780
Db  721   TTTACCCCTTTGGGGGTGATTGGTGATGTTACAACTCACACACTCTAGTAGGCCGCTGGT  780
Qy  781   GGCAGGAGGCGGTCGTTCGACCAGTCGTGCTCTGGAGAGTAGTACGCTTGCTGGAGGATGGAGTCAA  840
Db  781   GGCAGGAGGCGGTCGTTCGACCAGTCGTGCTCTGGAGAGTAGTACGCTTGCTGGAGGATGGAGTCAA  840
Qy  841   CTGGGCTACTGTGTTCGGTGTCACCTTTTTGTGTATGTCTGCTATCTTTGGCCTG  900
Db  841   CTGGGCTACTGTGTTCGGTGTCACCTTTTTGTGTATGTCTGCTATCTTTGGCCTG  900
Qy  901   TCCCTGTAGTGGGCGGCCGGGTCACTGACCCAGACACAAATACCACCAATCTGACCAATTG  960
Db  901   TCCCTGTAGTGGGCGGCCGGGTCACTGACCCAGACACAAATACCACCAATCTGACCAATTG  960
Qy  961   CTGCCAGCGTAATCAGTTATTGTTCCTTCCACTTGCCACTTGCCTACACGAGCCTCGGTTG  1020
Db  961   CTGCCAGCGTAATCAGTTATTGTTCCTTCCACTTGCCACTTGCCTACACGAGCCTCGGTTG  1020
Qy  1021  TGTGATCTGTGCGGACAGTGCTGGTGGTTCGGTTCCCGCCAATCCGTACATCCTCACACCCTTCCAA  1080
Db  1021  TGTGATCTGTGCGGACAGTGCTGGTGGTTCGGTTCCCGCCAATCCGTACATCCTCACACCCTTCCAA  1080
Qy  1081  TTGGGTGCTACCGGGACTGGGCGGCCGGGTTCTTCTTCTTGGTATTTGTTGTATGGGCGCTCTTGT  1140
Db  1081  TTGGGACTGGGCACGGACTGGGCGGCCGGGTTCTTCTTCTTGGTATTTGTTGTATGGGCGCTCTTGT  1140
Qy  1141  GACCTTGTGACGCCCTTGACATTGGTGCGTCGTATTAGTCGGTGACTG  1200
Db  1141  GACCTTGTGACGCCCTTGACATTGGTGCGTCGTATTAGTCGGTGACTG  1200
Qy  1201  GCTTTGTCAGGGACTGGCTTATTCACATAGACCTTGTACTTGTTACCTTGGA  1260
Db  1201  GCTTTGTCAGGGACTGGCTTATTCACATAGACCTTGTACTTGTTACCTTGGA  1260
Qy  1261  AGTGCCACTGGAATAGATCTCGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT  1320
```

```
Qy  2401 CTGCAGGCTACGTTATGCTGCCCTTTTAGGGTTTGTGCCCATGGCTGCGGGCTTGCCCCT 2460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 CTGCAGGCTACGTTATGCTGCCCTTTTAGGGTTTGTGCCCATGGCTGCGGGCTTGCCCCT 2460

Qy  2461 AACTTTCTTTGTTGCGCAGCAGCTGCTGCCCAACCAGATGGTGGTGTAACGACTGGGGTC 2520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 AACTTTCTTTGTTGCGCAGCAGCTGCTGCCCAACCAGATGGTGGTGTAACGACTGGGGTC 2520

Qy  2521 AGTGGCAGGGTAGTTTTGTGGGCGGCGTAACCGTGCCATAGCTCTGCTTGT 2580
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521 AGTGGCAGGGTAGTTTTGTGGGCGGCGTAACCGTGCCATAGCTCTGCTTGT 2580

Qy  2581 AGGTTCCTTGGCCTCTGGTAGCGCCTTTAACCCTCTTGGTTACGCCTGCTTCAGC 2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581 AGGTTCCTTGGCCTCTGGTAGCGCCTTTAACCCTCTTGGTTACGCCTGCTTCAGC 2640

Qy  2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT 2700
         |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641 TTTTGACACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT 2700

Qy  2701 GTCTCGTTTGGGCTTCTTGCTCACTTGTTACCTCGCTGTGTCTTAGTTAACTCCTATCT 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 GTCTCGTTTGGGCTTCTTGCTCACTTGTTACCTCGCTGTGTCTTAGTTAACTCCTATCT 2760

Qy  2761 TTGGCAACGTTGGGAGAATTGGTTTTCCCCGGTTGCGACATATGACGCGCTGGTGACTTTCCT 2820
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 TTGGCAACGTTGGGAGAATTGGTTTTCCCCGGTTGCGACATATGACACACGCTGGTGACTTTCCT 2820

Qy  2821 TGTGCTCGGTTTGTTTGTTTCCTTCTCATGTTAACATCCAGTGCAGCATCCAGTCGTTCTTCGTGTGTGTGTCA 2880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821 TGTGCTCGGTTTGTTTGTTTCCTTCTCATGTTAACATCCAGTGCAGCATCCAGTCGTTCTTCGTGTGTGTGTCA 2880

Qy  2881 CGTAGCTTCTTCTATGTTTAACATCCAGTGCAGCATCGTTCTTCGTGTGTGTTGGGACTGACTCTAGGGT 2940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881 CGTAGCTTCTTCTATGTTTAACATCCAGTGCAGCATCGTTCTTCGTGTGTGTTGGGACTGACTCTAGGGT 2940

Qy  2941 TAGGGCCCATAGAATGTTGGTGCGTCTCGGGAAGTGTCATGCTTGGTATTCTTCATTATGT 3000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 TAGGGCCCATAGAATGTTGGTGCGTCTCGGGAAGTGTCATGCTTGGTATTCTTCATTATGT 3000

Qy  3001 TCTTAAGTTTTTCCTCTCTAGTGTTTGGTGAGAATGGTGTGTTTTTCTATAAGCACTTGCA 3060
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 TCTTAAGTTTTTCCTCTCTAGTGTTTGGTGAGAATGGTGTGTTTTTCTATAAGCACTTGCA 3060

Qy  3061 TGGTGATGTCTTGCCTAATGATTTGCCTGGAAGCGCACCTTTTACGCTGCGACGTGTCTCTCAAGAGCCATTTTTCCC 3120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 TGGTGATGTCTTGCCTAATGATTTGCCTGGAAGCGCACCTTTTACGCTGCGACGTGTCTCTCAAGAGCCATTTTTCCC 3120

Qy  3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGAAGACGCTTGGCGTGTGGGGACAC 3180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGAAGACGCTTGGCGTGTGGGGACAC 3180

Qy  3181 GGTTGATGGTTTGCCCGTGTTGCCGCTCGTGTTGCCCGGACCTTGTTTTCGCAGGGGTTGGCTAT 3240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 GGTTGATGGTTTGCCCGTGTTGCCGCTCGTGTTGCCCGGACCTTGTTTTCGCAGGGGTTAGCTAT 3240

Qy  3241 GCCGCCAGATGGGTGGGCCATTACCGCACCTTTTACGCTGCGACGTGTCTCTGAACGTGG 3300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 GCCGCCAGATGGGTGGGCCATTACCGCACCTTTTACGCTGCGACGTGTCTCTGAACGTGG 3300

Qy  3301 CACGCTGTCAGCGATGGCAAGGGTCATGACTCGTATAGACCCCCGGACTTGGGGACAC 3360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 CACGCTGTCAGCGATGGCAAGGGTCATGACTCGTATAGACCCCCGGACTTGGGGACAC 3360

Qy  3361 TATCTTCAGATTAGGAATCTCGGCCACTAGCTACATGGGATTTGTTTGTGACAACGTGTT 3420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 TATCTTCAGATTAGGAATCTCGGCCACTAGCTACATGGGATTTGTTTGTGACAACGTGTT 3420

Qy  3421 GTATACTTGCTCACCATGGCGCCGGCGCAAGGGGCGCCCGGTTGGCTCATCCCACAGGCTCTATACA 3480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421 GTATACTTGCTCACCATGGCGCCGGCGCAAGGGGCGCCCGGTTGGCTCATCCCACAGGCTCCATACA 3480
```

```
Qy  3481 CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTCG 3540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481 CCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTCG 3540

Qy  3541 GTCCCTTACTCGGTCGTCTTGCGGGGAGACCAAGGGTATCTGTAACGACTGGGGTC 3600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541 GTCCCTTACTCGGTCGTCTTGCGGGGAGACCAAGGGTATCTGTAACGACTGGGGTC 3600

Qy  3601 ATTGGTTGAGGTCAACAAATCCGATGACCCCTTATTGGTGTGTGCGGGGCCCTCCCAT 3660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601 ATTGGTTGAGGTCAACAAATCCGATGACCCCTTATTGGTGTGTGCGGGGCCCTCCCAT 3660

Qy  3661 GGCTGTGTGCCAAGGGTTCTTCAGGTGCCCGATTCGTGTCGTCCTCCCGGCATGTTATTGG 3720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661 GGCTGTGTGCCAAGGGTTCTTCAGGTGCCCGATTCGTGTCGTCCTCCCGGCATGTTATTGG 3720

Qy  3721 GATGTTCACCGCTGCTAGAAATTCTGGCGGTTCAGTCAGTTCAGCCAGTACAAAACCTAC 3780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721 GATGTTCACCGCTGCTAGAAATTCTGGCGGTTCAGTCAGTTCAGCCAGTACAAAACCTAC 3780

Qy  3781 GGTGTGTGCTGGATACCATCCCCAGTACACAGCACATGCCACTTCTTGATACAAAACTCAAC 3840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781 GGTGTGTGCTGGATACCATCCCCAGTACACAGCACATGCCACTTCTTGATACAAAACTCAAC 3840

Qy  3841 TGTGCCTAACGTGATATTCAGTGACGAAGTATGAGTCTTGGTCCTAAATCCCAGTGT 3900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 TGTGCCTAACGTGATATTCAGTGACGAAGTATGAGTCTTGGTCCTAAATCCCAGTGT 3900

Qy  3901 CAAATTACCACTTTCTTACATGCCAAAGTAGAGGAGATGAGGTCAATCCAAATTG 3960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 CAAATTACCACTTTCTTACATGCCAAAGTAGAGGAGATGAGGTCAATCCAAATTG 3960

Qy  3961 GGCTACAACAGCATCAATGCCAAAGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGGCAT 4020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 GGCTACAACAGCATCAATGCCAAAGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGGCAT 4020

Qy  4021 CTATTTTAATGGCAAATGTACCAACACAGGGGCTTCACTTGATGTAATCATTGTGACGAATGCCATGC 4080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 CTATTTTAATGGCAAATGTACCAACACAGGGGCTTCACTTGATGTAATCATTGTGACGAATGCCATGC 4080

Qy  4081 GTACCTGACCGGAGCATGTCCCGGAACTGTTGGAAGGTCCTAACGAAGCTCCATCCAAAAA 4140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 GTACCTGACCGGAGCATGTCCCGGAACTGTTGGAAGGTCCTAACGAAGCTCCATCCAAAAA 4140

Qy  4141 TACCGATGCAACCACCGTGTTGGGCATTGGGCCGTGTTGGGCCACGCTAATCCTCACCACCATGC 4200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 TACCGATGCAACCACCGTGTTGGGCATTGGGCCGTGTTGGGCCACGCTAATCCTCACCACCATGC 4200

Qy  4201 TGTTTAGGGCTAGTAGCTGGAATAACAGTCGTCTCTCTTACTATAGGGG 4260
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201 TGTTTAGGGCTAGTAGCTGGAATAACAGTCGTCTCTCTTACTATAGGGG 4260

Qy  4261 CAACATAACTGAGATTCAATTAACCGATGAAGGGCACTATCCCCTTTCATGGGAAAAAGAT 4320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261 CAACATAACTGAGATTCAATTAACCGATGAAGGGCACTATCCCCTTTCATGGGAAAAAGAT 4320

Qy  4321 TAAGGAGGAAAATCTGAAGAAAGGGAGCACACCTTATCTTTGAGGCTACCAAAAAACACTG 4380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321 TAAGGAGGAAAATCTGAAGAAAGGGAGCACACCTTATCTTTGAGGCTACCAAAAAACACTG 4380

Qy  4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGTCGTCTCTTACTATAGGGG 4440

Qy  4441 ATGTGACATCTCAAAAATCCAAAAATCCCTGAGGGCGACTGTGTAGTAGTGCCACTGAATGCCTTGTG 4500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441 ATGTGACATCTCAAAAATCCAAAAATCCCTGAGGGCGACTGTGTAGTAGTGCCACTGAATGCCTTGTG 4500

Qy  4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTGGCACCTCATGGTAGAGGCAC 4560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTGGCACCTCATGGTAGAGGCAC 4560

Qy  4561 ATGCCATGTTGACCTTGACCCTGACCCTACTTTCACCATGGGTGTTCGTGTGTCGGGGGTTTCAGC 4620
```

```
Db 4561                                                                   4620
       |||||||||||||||||||||||||||||||||||||||||||||| |||||
Qy        ATGCCATGTTGACCTTGACCCTACTTTCACCATGGTGTTCGTGTGTGCGGGGTCTCAGC

Db 4621 AATAGTTAAAGGCCAGCGTAGGGGCCGCACAGGCGGCGTGGGAGAGCTGGCATATACTACTA 4680
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 4621 AATAGTTAAAGGCCAGCGTAGGGGCCGCACAGGCCGTGGGAGAGCTGGCATATACTACTA 4680

Db 4681 TGTAGACGGGAGTGTACCCCTTCGGGTATGGTATGGTTCCTGAATGCAACATTGTTGAGCCTT 4740
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 4681 TGTAGACGGGAGTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAGCCTT 4740

Db 4741 CGACGCAGCCAAGGCATGTATGGTTGTCATCAACAGAAGCTCAAACTATTCTGACAC 4800
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 4741 CGACGCAGCCAAGGCATGTATGGTTGTCATCAACAGAAGCTCAAACTATTCTGACAC 4800

Db 4801 CTATGCACCCCAACCTGGGTTACCTGCGATAGGAGCAAATTGGACGAGTGGGCTGATCT 4860
       |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 4801 CTATCGCACCCCAACCTGGGTTACCTGCGATAGGAGCAAATTGGACGAGTGGGCTGATCT 4860

Db 4861 CTTTTCTATGGTCAACCCCGAACCTTCATTTGTCAAATACTGCAAAAAGAACTGCTGACAA 4920
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 4861 CTTTTCTATGGTCAACCCCGAACCTTCATTTGTCAAATACTGCAAAAAGAACTGCTGACAA 4920

Db 4921 TTATGTTTTGTTGACTGCAGCCAACTACAACTGTGTCATCAGTATGCTATGCTGCTCC 4980
       ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 4921 TTATGTTTTGTTGACTGCAGCCAACTACAACTGTGTCATCAGTATGCTATGCTGCTCC 4980

Db 4981 CAATGACGCACCACGGTGGCAGGAGACCCGGCTTGTGGGGAAAAAACCTTGTGTGGGGTTCTGTG 5040
       ||||||||||||||||||||||||||||||||||||||||||||||||
Qy 4981 CAATGACGCACCACGGTGGCAGGAGACCCGGCTTGTGGGGAAAAAACCTTGTGTGGGGTTCTGTG 5040

Db 5041 GCGCTTGGACGGCGCTGAGCGCTGACGCGCTGGACGCGCTGGACGCGCTGACGCGCTGACGCGCTGACGCGCTGACGCGCTGACGCGCTGGACACACCAGATACCA 5100
       ||||||||||||||||||||||||||||||||||||||||||||||||
Qy 5041 GCGCTTGGACGGCGCTGACGCGCTGACACCAGATACCA 5100

Db 5101 AATGTGCTTCACTGAAGTCAATACTTCGGACAGCCGCACTCGCTCGTGTTGGCGGTTGGGAGT 5160
       |||||||||||||||||||||||||||||||||||||||||||||||||
Qy 5101 AATGTGCTTCACTGAAGTCAATACTTCGGACAGCCGCACTCGCTCGTGTTGGCGGTTGGGAGT 5160

Db 5161 GGCTATGGCTTATCTAGCCATTGACACTTTGGCGGCCCACTTGTGTGCCGGTTGCTGCGGTC 5220
       |||||||||||||||||||||||||||||||||||||||||||||||||
Qy 5161 GGCTATGGCTTATCTAGCCATTGACACTTTGGCGGCCCACTTGTGTGCCGGTTGCTGCGGTC 5220

Db 5221 TATTACATCAGTCCCGGTGCTACTGGTCGTCGCCTGGAGGCCATGGTTGACGAAGAGAAATCGT 5280
       ||||||||||||||||||||||||||||||||||||||||||||||||
Qy 5221 TATTACATCAGTCCCGGTGCTACTGGTCGTCGCCTGTGTGGTTGACGAAGAGAAATCGT 5280

Db 5281 GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGGTTGCTGCGAATTGACAAGCTGAA 5340
       |||||||||||||||||||||||||||||||||||||||||||||||||
Qy 5281 GGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCATGGTTGCTGCGAATCGATAAGCTGAA 5340

Db 5341 GAGTACAATCACCACAACCACAACCAGTCCTTCACATTGGAAACCGCCCCTTGAAAAAACTTAACAC 5400
       ||||||||||||||||||||||||||||||||||||||||||||||||
Qy 5341 GAGTACAATACCACAACCAGTCCTTCACATTGGAAACCGCCCCTTGAAAAAACTTAACAC 5400

Db 5401 CTTTCTTGGGGCTCATGCAGTCTACAATCCTTGCTATCATCAGAGTAGTATTGCTTGGTTTAGT 5460
       ||||||||||||||||||||||||||||||||||||||||||||||||
Qy 5401 CTTTCTTGGGGCTCATGCAGTCTACAATCCTTGCTATCATCAGAGTAGTATTGCTTGGTTTAGT 5460

Db 5461 CACTTTTACCTGACAATCCCTTTGCATCATGCGTGTTTTGCTTTTCATTGCGGGTATTACTAC 5520
       |||||||||||||||||||||||||||||||||||||||||||||||||
Qy 5461 CACTTTTACCTGACAATCCCTTTGCATCATGCGTGTTTTGCTTTTCATTGCGGGTATTACTAC 5520

Db 5521 CCCACTACCTCACAAGATCAAAGATCAAAAATGTTCCTTGCATTATTGGAGGCGCAATTGCGTCCAA 5580
       ||||||||||||||||||||||||||||||||||||||||||||||||
Qy 5521 CCCACTACCTCACAAGATCAAAGATCAAAAATGTTCCTTGCATTATTTGGAGGCGCAATTGCGTCCAA 5580

Db 5581 GCTTACAGACGCGCTAGAGGGCACTGGCGTTCATGATGGCCTCGGGGAACAGCTCT 5640
       |||||||||||||||||||||||||||||||||||||||||||||||
Qy 5581 GCTTACAGACGCGCTAGAGGGCACTGGCGTTCATGATGGCCTCGGGGAACAGCTCT 5640

Db 5641 TGGTACATGGACATCGGTGGGTTTGTCTTGACATGCTAGGCGGCTATGCTGCCGCCTC 5700
       |||||||||||||||||||||||||||||||||||||||||||
Qy 5641 TGGTACATGGACATCGGTGGGTTTGTCTTGACATGCTAGGCGGCTATGCTGCCGCCTC 5700
```

```
Qy 5641 TGGTACATGGACATCGGTGGGTTTGTCTTGACATGCTAGGCGGCTATGCTGCCGCCTC 5700

Qy 5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGCCCACTATGGATCAGCT 5760
       |||||||||||||||||||||||||||||||||||||||||||||||||
Db 5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGCCCACTATGGATCAGCT 5760

Qy 5761 TGCTGGTTAGTCTACTCCGGTTCAATCCGGCCGCAGGAGAGTGTGGGCGTCTTGTCAGC 5820
       ||||||||||||||||||||||||||||||||||||||||||||||
Db 5761 TGCTGGTTAGTACTCCGGTTCAATCCGGCCGCAGGAGAGTGTGGGCGTCTTGTCAGC 5820

Qy 5821 TTGTGCAATGTTGCTTGACAACAGCAGGGCCAGATCACTGCCCAACAGACTTCTTAC 5880
       |||||||||||||||||||||||||||||||||||||||||||||
Db 5821 TTGTGCAATGTTGCTTGACAACAGCAGGGCCAGATCACTGCCCAACAGACTTCTTAC 5880

Qy 5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940
       ||||||||||||||||||||||||||||||||||||||||||||||||
Db 5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940

Qy 5941 CAGGAGATACTGGCATTCTGGAGCATCCACCCCCTGGAGTGTCATATCAGCTTGCAT 6000
       ||||||||||||||||||||||||||||||||||||||||||||||
Db 5941 CAGGAAGATACTGGGCATTCTGGAGCATCCACCCCCTGGAGTGTCATATCAGCTTGCAT 6000

Qy 6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGGCCTCATTGCTTGGGGTCTAGAGAT 6060
       |||||||||||||||||||||||||||||||||||||||||||||||
Db 6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGGCCTCATTGCTTGGGGTCTAGAGAT 6060

Qy 6061 TTGGCAGTATGTGTGCAATTCTTGTGATTTGCTTTAAGTGTCCTTAAAGCTGGAGTTCA 6120
       |||||||||||||||||||||||||||||||||||||||||||||||
Db 6061 TTGGCAGTATGTGTGCAATTGTGCTTTAAGTGTCCTTAAAGCTGGAGTTCA 6120

Qy 6121 GAGCCATGGTAACATTCCTGGTTGTCCTTCACAGCTGCCAGAAGGGGTACAAGGGCCC 6180
       ||||||||||||||||||||||||||||||||||||||||||||||
Db 6121 GAGCCATGGTAACATTCCTGGTTGTCCTTCACAGCTGCCAGAAGGGGTACAAGGGCCC 6180

Qy 6181 CTGGATTGGATCAGGTATGCTCCAAGCACGCTGTCATGCGGTGCTGAACTCATCTTTTC 6240
       |||||||||||||||||||||||||||||||||||||||||||||||
Db 6181 CTGGATTGGATCAGGTATGCTCCAAGCACGCTGTCATGCGGTGCTGAACTCATCTTTTC 6240

Qy 6241 TGTTGAGAATGGTTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG 6300
       ||||||||||||||||||||||||||||||||||||||||||||||
Db 6241 TGTTGAGAATGGTTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG 6300

Qy 6301 AGGGGCTGTTCCAGTCAACGCTCAACCGTAGGCGTGTGGGTCGGCCGACCCAACTGATTG 6360
       |||||||||||||||||||||||||||||||||||||||||||||||
Db 6301 AGGGGCTGTTCCAGTCAACGCTCAACCGTAGGCGTGTGGGTCGGCCGACCCAACTGATTG 6360

Qy 6361 GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGGGAGAAATGGGAGA 6420
       |||||||||||||||||||||||||||||||||||||||||||||||
Db 6361 GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGGGAGAAATTGGGAGA 6420

Qy 6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCGTTTCACCCAGGTGCCCCCAAC 6480
       ||||||||||||||||||||||||||||||||||||||||||||||
Db 6421 TCACATTTTTGTTACAGCAGTATCCTCTCCAAATGTCGTTTCACCCAGGTGCCCCCAAC 6480

Qy 6481 CTTGAGAGCTGCAGTGCCGTGACGCGGCGGTTCAGTGTTATCTAGGTGAGCCCAA 6540
       |||||||||||||||||||||||||||||||||||||||||||||
Db 6481 CTTGAGAGCTGCAGTGCCGTGACGCGGCGGTTCAGTGTTATCTAGGTGAGCCCAA 6540

Qy 6541 AACTCCTTGGACGACATCGCTTGCTTGTCGGTCACACCTGTGTCGTGCGACTTAATTTGCGTGA 6600
       |||||||||||||||||||||||||||||||||||||||||||||||
Db 6541 AACTCCTTGGACGACATCGCTTGCTTGTCGGTCACACCTGTGTCGTGCGACTTAATTTGCGTGA 6600

Qy 6601 GCTTCCCTTCCCGGTTGACGGGTCACACCTGCGTGTTGCTGCATGCGAACTTAATTTGCGTGA 6660
       |||||||||||||||||||||||||||||||||||||||||||||||
Db 6601 GCTTCCCTTCCCGGTTGACGGGTCACACCTGCGTGTTGCTGCATGCGAACTTAATTTGCGTGA 6660

Qy 6661 TGCACTTGAGACAAATGACTGGTAATTCCACAAACAACACTCCTAGTGATGAAGCCGCAGT 6720
       ||||||||||||||||||||||||||||||||||||||||||||||
Db 6661 TGCACTTGAGACAAATGACTGGTAATTCCATAAACAACACTCCTAGTGATGAAGCCGCAGT 6720

Qy 6721 GTCCGCTCTTGTTTCAAACAGGAGTTGCGGCCGTACAAACCAATTGCTTGAGGCAATTTC 6780
       |||||||||||||||||||||||||||||||||||||||||||||||
Db 6721 GTCCGCTCTTGTTTCAAACAGGAGTTGCGGCCGTACAAACCAATTGCTTGAGGCAATTTC 6780
```

```
Qy   6781 AGCTGGCGTTGACACCACCAAAACTGCCAGCCCCCTCCATCGAAGAGGAGTAGTGTAAGAAA 6840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6781 AGCTGGGCGTTGACACCACCAAAACTGCCAGCCCCCTCCATCGAAGAGGAGTAGTGTAAGAA 6840

Qy   6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTCGTGCCTACCTTGCCTCCCCCTCCGAGATCCGTCCC 6900
          |||||||||||||||||||||||||||||||||||
Db   6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTCGTGCCTACCTTGCCTCCCCCTCCGAGATCCGTCCC 6900

Qy   6901 AGGAGTGTCATGTCCTGAAAGCCTGCGACGAAGTGACCCGGTTAGAAGGTCCTTCAAACCT 6960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6901 AGGAGTGTCATGTCCTGAAAGCCTGCGACGAAGTGACCCGGTTAGAAGGTCCTTCAAACCT 6960

Qy   6961 CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGCCGATGCCCGTTGGGAGCGGGGTGA 7020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6961 CCCTTCTTCACCACCTGTTCTACAGTTGGCCATGCCGATGCCCGTTGGGGAGCGGGGTGA 7020

Qy   7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAACAGGCGAGAGGCCCTGATGA 7080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAACAGGCGAGAGGCCCTGATGA 7080

Qy   7081 TTTACCCAGTTACCCTCCCAAAAAGAGGAGTCTCTGAATGGTCAGACGAAAGTTGGTCGAC 7140
          |||||||||||||||||||||||     |||||||||||||||||||||||||||
Db   7081 TTTACCCAGTTACCCTCCCAAAAAGAGGAGTCTCTGAATGGTCAGACGAAAGTTGGTCAAC 7140

Qy   7141 GGCTACAACCGCTTCAGCTACGTTACGTGGCCCCCGGTACCCTAAGATACGGGGAAAGGA 7200
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7141 GACTACAACCGCTTCCGCGTCTACTGGCCCCCCGGTACCCTAAGATACGGGGAAAGGA 7200

Qy   7201 TTCCACTCAGTCAGCCGCCCCGCCAAAACGGCCTACAAAAAGAAGAGTTGGGAAAGAGTGAGTT 7260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7201 TTCCACTCAGTCAGCCGCCCCGCCAAAACGGCCTACAAAAAGAAGAGTTGGGAAAGAGTGAGTT 7260

Qy   7261 TTCGTGCAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7261 TTCGTGCAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320

Qy   7321 TCTGTCTGCAACTCGGGCCATCACTAGTGGTTTCCTCAAACAAAGATCATTGGTGTATGT 7380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7321 TCTGTCTGCAACTCGGGCCATCACTAGTGGTTTCCTCAAACAAAGATCATTGGTGTATGT 7380

Qy   7381 GACTGAGCGCGGGGGATGCGGACGCTTAGAAGCAAGCGAGTTTAAGAAAACAAACCTCT 7440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7381 GACTGAGCGCGGGGGATGCGGACGCTTAGAAGCAAGCGAGTTTAAGAAAACAAACCTCT 7440

Qy   7441 GTTCCCCCATCATACCACAAGCAGTGAGATTGGCTAAGGAAAAGCTTCAAAGTTGT 7500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7441 GTTCCCCCATCATACCACAAGCAGTGAGATTGGCTAAGGAAAAGCTTCAAAGTTGT 7500

Qy   7501 CGGTGTCATGTGGGACTATGATGAAGTAGCAGCTCACACGCCCTCACGCGCCGCAGGT 7560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7501 CGGTGTCATGTGGGACTATGATGAAGTAGCAGCTCACACGCCCTCACGCGCCGCAGGT 7560

Qy   7561 CCACATCACTGGCCTTCGGGGCGCTGATGTTCGTTCTTGGAGCAGCCCGCCAAGGCTGTCT 7620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7561 CCACATCACTGGCCTTCGGGGCGCTGATGTTCGTTCTTGGAGCAGCCCGCCAAGGCTGTCT 7620

Qy   7621 GGACTTGCAGAAGTGTGTCGAGCGCAGGTGAGCTGAGTCATTATCGGCAAACTGTGAT 7680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7621 GGACTTGCAGAAGTGTGTCGAGCGCAGGTGAGCTGAGTCATTATCGGCAAACTGTGAT 7680

Qy   7681 AGTTCCAAAGAAGGAGGAGTCTTCGTGAAGACACCCCAGAAACCAACAAAGAAACCCCCCAAG 7740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7681 AGTTCCAAAGAAGGAGGAGTCTTCGTGAAGACACCCCAGAAACCAACAAAGAAACCCCCCAAG 7740

Qy   7741 GCTTATCCTCGTACCCCCTTGAAATGAGGATGTGTTGAGAAGATGCTACGGGTTTGTAGATCCACGTCAGGT 7800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7741 GCTTATCTCGTACCCCCACCTTGAAATGAGGATGTGTTGAGAAGATGCTACGGGTTTGTAGATCCACGTCAGGT 7800

Qy   7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTACGGGTTTGTTGACCCCACGTAC 7860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTACGGGTTTGTTGACCCCACGTAC 7860
```

```
Qy   7861 CCGTGTCAAGCGTCGTTGTCGATGTGGTCACCCGATGCAGTCGGAGCCACATGCGATAC 7920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7861 CCGTGTCAAGCGTCGTTGTCGATGTGGTCACCCGATGCAGTCGGAGCCACATGCGATAC 7920

Qy   7921 AGTGTGTTTGACAGTACCATCACACCCGAGATATCATGGTGAGACAGACATCTACTC 7980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7921 AGTGTGTTTGACAGTACCATCACACCCGAGATATCATGGTGAGACAGACATCTACTC 7980

Qy   7981 AGCAGCTAAACTCAGTGACCAGAGCTGGCATTCACACCATTCACACCATTGCGAGGCAGTTATA 8040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7981 AGCAGCTAAACTCAGTGACCAGAGCTGGCATTCACACCATTCACACCATTGCGAGGCAGTTATA 8040

Qy   8041 CGCTGGAGGACCGATGATCGCTTATGATGGCCCGAGAGATCGGATATCGGATAGTTGTAGGTC 8100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8041 CGCTGGAGGACCGATGATCGCTTATGATGGCCCGAGAGATCGGATATCGGATAGTTGTAGGTC 8100

Qy   8101 TTCCGGCGTCTATACTACCTCAAGTTCCAACAGTTGCGGCGGATGATTGCAC 8160
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8101 TTCCGGCGTCTATACTACCTCAAGTTCCAACAGTTGCGGCGGATGATTGCAC 8160

Qy   8161 TGCAGCCGAACAGGCTGCAGATGAAGAACCCTCGCTTCCTTATTGCGGCGATGATTGCTAG 8220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8161 TGCAGCCCAACAGGCTGCAGATGAAGAACCCTCGCTTCCTTATTGCGGCGATGATTGCTAG 8220

Qy   8221 CGTAATTTGGAAGAGCGCCGGAGCAGATGCAGACAAGCAATGCAACAAGCAAGCAAGGA 8280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8221 CGTAATTTGGAAGAGCGCCGGAGCAGATGCAGACAAGCAATGCAACAAGCAAGCAAGGA 8280

Qy   8281 CTGGATGAAGGTGATGGGTGCACCACAAGATTGTGTGCCTCAAAATACAGTTTGGA 8340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8281 CTGGATGAAGGTGATGGGTGCACCACAAGATTGTGTGCCTCAAAATACAGTTTGGA 8340

Qy   8341 AGAATTAACATCATGCTCATCAAATGTTACCCTGGAATTACCAAAAGTGGCAAGCCTTA 8400
          ||||||||||||||||||||||||||||||  |||||||||||||||||||||||||
Db   8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400

Qy   8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGTGCTCTGCCGAGGGTCTGGG 8460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGTGCTCTGCCGAGGGTCTGGG 8460

Qy   8461 ATACAACCCCAGTGCTGCTGGTGATTGGGTATCTAATACATCAATGACCATGTTTGTGGGT 8520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8461 ATACAACCCCAGTGCTGCTGGTGATTGGGTATCTAATACATCAATGACCATGTTTGTGGGT 8520

Qy   8521 TAGCCGTGTTGGCTGTGGCTCTTCATTCATGGATGGGAAAAATTATACGGTGTAGAAGATCTCCCGA 8580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8521 TAGCCGTGTTGGCTGTGGCTCTTCATTCATGGATGGGAAAAATTATACGGTGTAGAAGATCTCCCGA 8580

Qy   8581 GACTGACCTTTGACTGTGTATGGGAAAAAATTATACGGTGTAGAAGATCTGCCCAG 8640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8581 GACTGACCTTTGACTGTGTATGGGAAAAAATTATACGGTGTAGAAGATCTGCCCAG 8640

Qy   8641 CATCATTGCTGGTGTGCACGTATTGAGGCTTTCTCGGTGGCGCTACACCAACGCTGA 8700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8641 CATCATTGCTGGTGTGCACGTATTGAGGCTTTCTCGGTGGCGCTACACCAACGCTGA 8700

Qy   8701 GATCCTCAGAGTTTCCAATCACTACGACATGAACACAGACATGACCATGCCCCCTGCGAGCCTGGCG 8760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8701 GATCCTCAGAGTTTCCAATCACTACGACATGAACACAGACATGACCATGCCCCCTGCGAGCCTGGCG 8760

Qy   8761 AAAGAAAGCCAGGGCGTCTTCGTGAGACACGCCCCAGAGGGCGGCGAGCACGCAAATT 8820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8761 AAAGAAAGCCAGGGCGTCTTCGTGAGACACGCCCCAGAGGGCGGCGAGCACGCAAATT 8820

Qy   8821 GGCTTCGCTTCCTTCTCTGGCATGCTACATCTAGACCCTCTACAGATTTGGATAAGACGAG 8880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8821 GGCTTCGCTTCCTTCTCTGGCATGCTACATCTAGACCCTCTACAGATTTGGATAAGACGAG 8880

Qy   8881 CGTTGGGCTCGGTACACCACTTTACTCCCCGGAGGGGATGTGTT 8940
          |||||||||||||||||||||||||||||||||||||||||||
Db   8881 CGTTGGGCTCGGTACACCACTTTACTCCCCGGAGGGGATGTGTT 8940

Qy   8941 TATTACACCAGAGAAGAATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTT 9000
Db   8941 ...
```

```
Db   8941 ||||||||  ||||||||||||||||||| |||||||||||||||| |||||||||||||
          TGTTACACCAGAGAGAAGATTGCAGAAGTTTCTGTGAAGTATTGGCTGTCATTGTTTT 9000

Qy    196 TGGGTTCGGTGGTGGTGGCGCTTAGCAGCTTCCACGCCCACCACCTCCCAGATAGAGC 255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11 TGGGTTCGGTGGTGGTGGCGCTTAGCAGCTTCCACGCCCACCACCTCCCAGATAGAGC 70

Qy   9001 TGCCCTAGGGGCTCATTGCTGTTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA 9060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9001 TGCCCTAGGGGCTCATTGCTGTTGGACTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA 9060

Qy    256 GGCGGCACTGTAGGGAAGAACCGGGACCGGGACCGGTCACTACCAAGACGGACGACCTCTTTTTGA 315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     71 GGCGGCACTGTAGGGAAGAACCGGGACCGGGACCGGTCACTACCAAGACGCAGACCTCTTTTTGA 130

Qy   9061 CTAACAG----TTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGAGACCCC 9116
          ||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   9061 TTAACAGTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGAGACCCC 9120

Qy    316 GTATCAGCGCCTCCGGAGTAGTTGGCAAGCCCACCTATATGTTGGGATGGTTGGGGT 375
          ||||||||||||||||||||||||||||||| :|||||||||||||||||||||||
Db    131 GTATCACGCCTCCGGAGTAGTTGGCAAGCCCACCTAYATGTTGGGATGGTTGGGGT 190

Qy   9117 GGGCTTAACGACCCCGC 9133
          ||||||||||||||
Db   9121 GGGCTTAACGACCCCGC 9137

Qy    376 TAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGGGAGGGATCTGGAGTCTCGTAG 435
          |||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db    191 TAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGGGAGGGATCTGGAGTCTCGTAG 250

Qy    436 ACCGTAGCACATGCCTGTATTTCTACTCAAACAAGTCCTGTACCTGCGCCCAGAACGCG 495
          ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db    251 ACCGTAGCACATGCCTGTATTTCTACTCAAACAAGTCCTGTACCTGCRCCCAGAACGCG 310

Qy    496 CAAGAACAAGCAGCAGGCTTCATATCCTGTGTCCATTAAAACATCTGTGAAAGGGG 555
          |||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db    311 CAAGAACAAGCAGACAGGCTTCATATCCTGTGTCCATTAAAAACATCTGTGAAAGGGG 370

Qy    556 ACAACGAGCAAAGTCCAGCGCGATGCTCGGCCTCGGCTGCTTGCTCCAGTCATGGTTGGG 615
          ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db    371 ACAACGAGCAARGGCAAAGTCCAGCGCGATGCTCGGCCTCGGCTGCTTGCTCCAGTCATGGTTGGG 430

Qy    616 TATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTGCCAGTCCATGGTTGGGACG 675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    431 TATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTRCCAGCTCCATGGTTGGGACG 490

Qy    676 CCAAGACCCTCGCCATAGTCTCGCAATCTTGGAATCTTCTGGATTACCCTTTGGGGTG 735
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    491 CCAAGACCCTCGCCATAAGTCTCGCAATCTTGGAATCTTCTGGATTACCCTTTGGGGTG 550

Qy    736 GATTGGTGATGTTACAACTCACACACCTCAGTAGGGCCCGGTGGTGGCAGGAGCGGTCGT 795
          ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db    551 GATTGGTGATGTTACAACTCACACACCTCAGTAGGGCCCGGTGGTGGCAGGAGCGGTCGT 610

Qy    796 TCGACCAGTCTGCCAGATAGTACGCTTGCTGGGAGGATGGAGTCAACTGGGCTACTGGTTG 855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    611 TCGACCAGTCTGCCAGATAGTACGCTTGCTGGGAGGATGGAGTCAACTGGGCTACTGGTTG 670

Qy    856 GTTCGGTGTCCACCTTTTTGTGTATGTCTGCTCATCTTTGGCCGTCTCCTGTAGTGGGGC 915
          ||||||||||||||||||||:|:|||||||||||||||||||||||||||||||||||
Db    671 GTTCGGTGTCCACCTTTTTGTGTATGTCTGCTATYTTGGCCGTCTCCTGTAGTGGGGC 730

Qy    916 GCGGGTCACTGACCCAGACACAAATACCACAATCCTGACCAATTGCTGCCAGCGTAATCA 975
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    731 GCGGGTCACTGACCCAGACACAAATACCACAATCCTGACCAATTGCTGCCAGCGTAATCA 790

Qy    976 GGTTATCTATTGTTCTCCTTCCACTTGCCTACACGAGCCTGGTTGTGTGTGCGGA 1035
          |||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db    791 GGTTATCTAYTGTTCTCCTTCCACTTGCCTACACGAGCCTGGTTGTGTGTGCGYGGA 850

Qy   1036 CGAGTGCTGGGGTTCCCGCCAATCCGTACATCTCACACCCTTCCAATTGGACTGCACCGGA 1095
          ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db    851 CGAGTGCTGGGGTTCCCGCCAATCCRTACATCTCACACCCTTCCAATTGGACTGCACCGGA 910

Qy   1096 CTCCTTCTTGCTGACCACATTGATTTTGTATGGGCGCTCTTGTGACCTGTGACGCGCCCT 1155
          ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db    911 CTCCTTCTTGGCTGACCACATTGATTTTGTATGGGCGCTCTTGTGACCTGTGACGCGCCCT 970

Qy   1156 TGACATTGGTGAGTTGTGTGCGTGTATTAGTCGTGTGTATTGCTTGTCAGGCACTG 1215
          ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db    971 TGACATTGGTGAGTTGTGTGCGTGTGTATTAGTCGTGTGTATTGCTTGTCAGGCACTG 1030

Qy   1216 GCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGAAT 1275
          |||||||||||||||||||:||||||||||:||||||||||||||||
Db   1031 GCTTATTCACATAGACCTCAATGAAACTCAATGAAACTGGTACTTGTTACCTGGAAYGCCTACTGGAAT 1090

Qy   1276 AGATCCTGGGTTCCTAGGGTTTTATCGGGTGGATGGCCGCAAGGTCGAAGGCTGCATCTT 1335
```

RESULT 10
US-08-469-260A-11
; Sequence 11, Application US/08469260A
; Patent No. 6451578

GENERAL INFORMATION:
  APPLICANT: JOHN N. SIMONS
  APPLICANT: TAMI J. PILOT-MATIAS
  APPLICANT: GEORGE J. DAWSON
  APPLICANT: GEORGE G. SCHLAUDER
  APPLICANT: SURESH M. DESAI
  APPLICANT: THOMAS P. LEARY
  APPLICANT: ANTHONY SCOTT MUERHOFF
  APPLICANT: JAMES C. ERKER
  APPLICANT: SHERI L. BUIJK
  APPLICANT: ISA K. MUSHAHWAR
  TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
  TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
  NUMBER OF SEQUENCES: 716
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
    STREET: 100 ABBOTT PARK ROAD
    CITY: ABBOTT PARK
    STATE: IL
    COUNTRY: USA
    ZIP: 60064-3500
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/469,260A
    FILING DATE:
  CLASSIFICATION:
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US/08/424,550
    FILING DATE:
  ATTORNEY/AGENT INFORMATION:
    NAME: POREMBSKI, PRISCILLA E.
    REGISTRATION NUMBER: 33,207
    REFERENCE/DOCKET NUMBER: 5527.PC.01
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 708-937-6365
    TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 11:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 8912 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
US-08-469-260A-11

Query Match 92.5%; Score 8692.6; DB 3; Length 8912;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 8781; Conservative 56; Mismatches 52; Indels 10; Gaps 8;

(Sequence alignment output — two columns of paired database (Db) and query (Qy) nucleotide sequences with alignment bars.)

Right column:

```
Db 2169 AGTAGCTCMTTTGCTTTTGATTTTCTTCATCTGTGCTATCTCCGCTGCAGGCTACGTTA 2228
Qy 2416 TGCTGCCCTTTTAGGGTTTGTTGCTGCCCATGCGTGCGCGTGCGGGCTGCCCTAACCTTCTTTGTTGC 2475
Db 2229 TGCTGCCCTTTTAGGGTTTGTTGCTGCCCATGCGTGCGGGGCTGCCCTAACCTTCTTTGTTGC 2288
Qy 2476 AGCAGCTGCTGCCAACCAGATTATGACTGCTAGTGCGACTGCTGCTGGGGTTAGT 2535
Db 2289 AGCAGCTGCTGCCAACCAGATTATGACTGCTAGTGCGACTGCTGCTGGCAGGGTTAGT 2348
Qy 2536 TTTGTGGGCCGGCCGTAACCGTGGTCACCGCATAGCTCTGCTTGTAGGTCCTTGGCCTCT 2595
Db 2349 TTTGTGGGCCGGCCGTGACCGTGGTCA-CGCATAGCTCTGCTTGTAGGTCCTTGGCCTCT 2407
Qy 2596 GGTAGCGCGCTT-TAACCCTTCTGCATTTGGTTACGGCCTTTGATACCGAGA 2654
Db 2408 GGTAGCGCGTTTYTAACCCCTCTTGSSTKACGCCCTTGCTT-AGCTTTTGACACCGAGA 2466
Qy 2655 TAATTGGAGGGCTGACAAATACCACCTGTAGTAGCATTAGTTGTCATGTCTCGTTTGGCT 2714
Db 2467 TAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCATGTCTCGTTTGGCT 2526
Qy 2715 TCTTTGCTCACTTGTTACCTCGTCTGTGTCTTTAGTTAACTCCTATCTTTGGCAACGTTGGG 2774
Db 2527 TCTTTGCTCACTTGTTACCTCGTCTGTGTCTTTAGTTAACTCCTATCTTTGGCAACGTTGGG 2586
Qy 2775 AGAATTGGTTTTGGAACGTTACACTAAGACCGAGAGGTTTTTCCTTGGTCTGGTTTGTT 2834
Db 2587 AGAATTGGTTTTGGAACGTTACACTAAGACCGAGAGGTTTCTCCTTGYGCTGGTTTGTT 2646
Qy 2835 TCCCCGGTGCGACATATGACGCGCTGGTGACTTTCTGTGTGTCACGTAGCTCTTCTAT 2894
Db 2647 TCCCCGGTGCGACATATGACGTGCTGGTGACWTTCTGTGTGTCACGTAGCTCTTCTAT 2706
Qy 2895 GTTTAAACATCCAGTGCAGCCATCGTCTTTTGGGACTGACTCTAGGGGTTAGGGCCCATAGAA 2954
Db 2707 GTTTAACATCCAGTGCAGCAYMGTTCTTTGGGACTGACTCTAGGGGTTAGGGCCCATAGAA 2766
Qy 2955 TGTTTGGTGCGTCTCGGAAAGTGTCATGCTTGGTATTCATTATGTCTTAAGTTTTTCC 3014
Db 2767 TGTTTGGTGCGTCTCGGAAAGTGTCATGCTTGGTATTCATTATGTCTTAAGTTTTTCC 2826
Qy 3015 TCTTTAGTGTTTGGTGAGAATGGTGTGTTTTCTATAAAGCACTTGCATGGTGATGTCTTGC 3074
Db 2827 TCTTTAGTGTTTGGTGAGAATGGTGTGTTTTCTAKAAGCACTTGCATGGTGATGTCTTGC 2886
Qy 3075 CTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCATTTTCCCCTTTTGAAGGCAAGG 3134
Db 2887 CTAATGATTTTGCCTCGAAACTCGAAACTACCATTGCAAGAGCCATTTTCCCTTTTGAAGGCAAGG 2946
Qy 3135 CAAGGGTCTATAGGAATGAAGGAAGAGACCTTGGCGTGTGGGACACGGTTGGTTGC 3194
Db 2947 CAAGGGTCTATAGGAATGAAGGAAGAGACCTTGGSKKGTGGGGACACGGTTGATGGTTTGS 3006
Qy 3195 CCGTTGTTGCGGTCTCGGGACCTTGTTTCGGCGAGGGTTGGCTATGCGCGGCGAGATGGGT 3254
Db 3007 SCGTTGTBGCGGTCTCGGGACCTTGTTTCGGCGAGGGTTAGCTATGCGCGCGCCAGATGGGT 3066
Qy 3255 GGGCCATTACCGCACCTTTTACGCTGACAGTGTCTCTGAACGTGGCACGCGTGTCAGCGA 3314
Db 3067 GGGCCATTACCGCACCTTTTACGCTGCAGTGTCTCTGAACGTGGCACGCGTGTCAGCGA 3126
Qy 3315 TGGCAGTGGTCATGACTGGTATAGACCCCGGACTGGAACTATCTTCAGATTAG 3374
Db 3127 TGGCAGTGGTCATGACTGGTATAGACCCCGGACTGGAACTATCTTCAGATTAG 3186
Qy 3375 GATCTCTGCCACTAGCTACATGGGATTTGTTTGTGACAACGTGTTGTATACTGCTCACC 3434
Db 3187 GATCTCTGCCACTAGCTACATGGGATTTGTTTGTGAATACTGCTCACC 3246
Qy 3435 ATGGCAGCAAGGGGCGCCGGTTGCGTCATCCCAACGGGCCCAATAACCCGTTG 3494
Db 3247 ATGGCAGCAACGGGCGCCGGTTGCGTCATCCCAACGGGCCCAATAACCCGTTG 3306
```

Left column:

```
Db 1091 AGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGTCGAGGCTGTCATCTT 1150
Qy 1336 CTTGACCAAACTGGCTTCACAAGTACCATGACGCTATTGCGACTATGTTAGCAGTGTACA 1395
Db 1151 CTTGACCAAACTGGCTTCACAAGTACCATGACGCTATTGCGACTATGTTAGCAGTGTACA 1210
Qy 1396 CTACCTGGCGGTTGGCGGCTCTGATCTACTATGCCTCTCGGGGCAAGTGGTATCAGTTGCT 1455
Db 1211 CTACCTGGCGGTTGGCGCTCTGATCTACTACTAYGCCTCTCGGGGCAAGTGGTATCAGTTGCT 1270
Qy 1456 CCTAGGCGCTTATGCTTTTACATAGAAGGACCTTCTGGAAACCCATCAGGGTGCCCACTGG 1515
Db 1271 CCTAGGCGCTTAYGCTTTTACATAGAAGGACCTTCTGGAAACCCYATCAGGGTGCCCACTGG 1330
Qy 1516 ATGCTCAATAGCTGAGTTTGCTCGCCTTGATGATACCATGTCCTTGCCACTCTTATTT 1575
Db 1331 ATGCTCAATAGCTGAGTTTGCTCGCCTTGATGATACCATGTCCTTGCCACTCTTATTT 1390
Qy 1576 GAGTGAGAAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGACCAGGCCTATCACTCT 1635
Db 1391 GAGTGAGAAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGACCAGGCCTGTCACTCT 1450
Qy 1636 AGAGTATAACAACTCCATATCTTGGTACCCCATACAATCCCTGGTTGCGAGGGATGTAT 1695
Db 1451 AGAGTATAABAACTCCATATCTTGGTACACAATCCCTGGTTGCGAGGGGATGTAT 1510
Qy 1696 GGTTAAATTCAAAAATAACACATGGGGTTGCTGCCGTATTCGCGAATGTCCATCGTACTG 1755
Db 1511 GGTTAAATTCAAAAATAACACATGGGGTTGCTGCCG-WWTCGCAATGTCCATCGTACTG 1569
Qy 1756 CACTATGGGCACTGATGCAGTGTGGAACGAACACTCCGCAACACTTACGAAGCATGCGGTGT 1815
Db 1570 CACTATGGGCACTGATGCAGTGTGGAASSACAGTCGCACACTTACGAAGCATGCGGTGT 1629
Qy 1816 AACACCACATGGCTAACACCGCATGGCACAAACGGCTCAGCCCTGAAATTGGCTATATTACA 1875
Db 1630 AACACCACATGGCTAACACCGCATGGCACAACGGCTCAGCCCTGAAATTGGCTATATTACA 1689
Qy 1876 ATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTCAGGCCATTGTATTT 1935
Db 1690 ATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTCAGGCCACTGTGTATTT 1749
Qy 1936 TGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAATTCCACTCTCCTACC 1995
Db 1750 TGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAATTCCACTCTCCTACC 1809
Qy 1996 ACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCTGTTGGTACGTTCTTGGTTACA 2055
Db 1810 ACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCTGTTGGTACGTTCTTGGTTACA 1869
Qy 2056 GGTTCCGCAAGGGGTTTTACAGTGATGTGAAAGACCACAGGATTGATCACCAAAGA 2115
Db 1870 GGTTCCGCAA-GGTTTTACAGTGATGTGAAAGACCACAGGATTGATCACCAAAGA 1928
Qy 2116 CAAAGCCTGGAAAAATTATCAGTCTTATATTCCGGCTTTGTGTCTCTTTACGGG 2175
Db 1929 CAAAGCCTGGAAAAATTATCAGYTCTTATATTCCGGCTGTCTTGTTTGTCTCTTTACGGG 1988
Qy 2176 AGTTACCACCAAGGGCCGTGGTGTGCTAATTCTGTTGGGCAGCAAGTGATCTTAT 2235
Db 1989 AGTTACCACCAAGGGCCGTGGTGTGCTAATTCTGTTGTGGGCAGCAAGTGATCTTAT 2048
Qy 2236 TTTTAGCCTACCTTCGTCCCTTTGTTTTTGGGGCGGCTTCTCGGGTTACCCTTTGCG 2295
Db 2049 TTTTAGCCTACCTTCGTCCCTTTGTTTTTGGGGCGGCTTCTGTTACMCTTTGCG 2108
Qy 2296 TCCGTGTGCTCCATCCCAGTGTATCTCCAAGCTGGCTGGGATGTTTTGTTCTAAGCTCA 2355
Db 2109 TCCGTGTGCTCCATCCCAGTGTATCTCCAAGCTGGCTGGGATGTTTTGTTCTAAAGCTCA 2168
Qy 2356 AGTAGCTCMTTTTGCTTTTGATTTTCTTCATCTGTGCTATCTCCGCTGCAGGCTACGTTA 2415
```

This page consists of a two-block nucleotide sequence alignment between a query sequence (Qy) and a database sequence (Db), with vertical bars indicating matched positions. Each line is labeled "Qy" or "Db" with flanking coordinate numbers.

Left-hand alignment block (coordinate ranges approximately 3307–5714):

Label	Start	...	End
Qy	3495	ACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGAGCTGGGTCCTTACTCGGT	3554
Db	3307	ACGCGGCTAATGACCAGGACATCTATCAACCACCATGTGAGCTGGGTCCTTACTCGGT	3366
Qy	3555	GCTCTTGCGGGGAGACCAAGGGGTATCTGTAACAGACTGGGTCATTGGTTGAGGTCA	3614
Db	3367	GCTCTTGCGGGGAGACCAAGGGGTATCTGTAACAGACTGGGTCATTGGTTGAGGTCA	3426
Qy	3615	ACAAATCCGATGACCCTTATTGGTTGCGGGGCCCTTCCCATGGCTGTTGCCCAAGG	3674
Db	3427	ACAAATCCGATGACCCTTATTGGTTGCGGGGCCCTTCCCATGGCTGTTGCCCAAGG	3486
Qy	3675	GTTCTTCAGGTGCCCCGATTCTGTGCTCTCCGGGCATGTGTATTGGGATGTTCACCGCTG	3734
Db	3487	GTTCTTCAGGTGCCCCGATTCTGTGCTCTCCGGGCATGTGTATTGGGATGTTCACCGCTG	3546
Qy	3735	CTAGAAATTCTGGCGTCAGTGAGGTTAGGGTTAGCCGTGGTCCGGTGTGTGTCGGAT	3794
Db	3547	CTAGAAATTCTGGCGTCAGTGAGGTTAGGGTTAGCCGTGGTCCGGTGTGTGTCGGAT	3606
Qy	3795	ACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTACTGTGCCTAACGAGT	3854
Db	3607	ACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTACTGTGCCTAACGAGT	3666
Qy	3855	ATTCAGTGCAAATTTAATTGCCCCCACTGCAGCCGCAAGTCAACCAAATTACCACTTT	3914
Db	3667	ATTCAGTGCAAATTTAATTGCCCCCACTGCAGCCGCAAGTCAACCAAATTACCACTTT	3726
Qy	3915	CTTACATGCAGGAGAGTATGAGGTCTTGGTCTCCAAATCCCAGTGTGGCTACAACGCAT	3974
Db	3727	CTTACATGCAGGRGAAGYATGAGGTCTTGGTCTCCAAATCCCAGTGTGGCTACAACAGCAT	3786
Qy	3975	CAATGCCAAAGTACTATGCACGCGACGTACGGCGTGAATCCAAATTGCTATTTTAATGGCA	4034
Db	3787	CAATGCCAAAGTACTATGCACGCGACGTACGGCGTGAATCCAAATTGCTATTTTAATGGCA	3846
Qy	4035	AATGTACCAACACACAGGGGCTTCACTTACGTACAGCACATATGGCATGTACCTGACCTGGAG	4094
Db	3847	AATGTACCAACACACAGGGGCTTCACTTACGTACAGCACATATGGCATGTACCTGACCCGGAC	3906
Qy	4095	CATGTTCCCGGAACTATGATGATGTAATCATTGTGACGAATGCCATGCCATGCAACCA	4154
Db	3907	GATGTTTCCCGGAACTATGATGATGTAATCATTGTGACGAATGCCATGCCATGCGAACCA	3966
Qy	4155	CCGTGTTGGGCATTGGAAAGTCCTAACCGAAGCTCCATCCAAAAATGTTAGGCTAGTGG	4214
Db	3967	CCGTGTTGGGCATTGGAAAGTCCTAACCGAAGCTCCATCCAAAAATGTTAGGCTAGTGG	4026
Qy	4215	TTCTTGCCACGGCTACCCCCGGAGTAATAACAGCTGTTCTTACTACCACACATGCCAACATGTGACATCTCAA	4274
Db	4027	ACGAGTTAGCTCGAAAGGGAATAACAGCTGTTCTTACTACCACCGGATGTGACATCTCAA	4266
Qy	4275	TTCAAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAGATTAAGGAGGAAAATC	4334
Db	4087	TTCAAATTAAACYGATGAAGGCACTATCCCCTTTCATGGAAAAAGATTAAGGAGGAAAATC	4146
Qy	4335	TGAAGAAAGGGAGACACACCTTATCTTTGAGGCTACCAACATGCCAACATAACACTGTGATGAGCTTGCTA	4394
Db	4147	TGAAGAAAGGGAGACACACCTTATCTTTGAGGCTACCAACATAACACTGTGATGAGCTTGCTA	4206
Qy	4395	ACAGTTAGCTCGAAAGGGAATAACAGCTGTTCTTTACTACTAGGGATGTGACATCTCAA	4454
Db	4207	ACGAGTTAGCTCGAAAGGGGAATAACAGCTGTTCTTACTACTAGGGATGTGACATCTCAA	4266
Qy	4455	AAATCCCTGAGGGCGGACTGTGTGTAGTGTTGCCACTGATGCTTGTGTCTGTGTACAGGGTACACTG	4514
Db	4267	AAATGCCTGAGGGCGGACTGTGTAGTGTTGCCACTGATGCTTGTGTCTGTGTACAGGGTACACTG	4326
Qy	4515	CTGACTTTGATTCCGTGTATCGCGTAATGACTGCAGCCTCATGGTGAGAAGGCACATGCCATGTTGACC	4574
Db	4327	GTGACTTTGATTCCGTGTAATGACTGCAGCCTCATGGTGAGAAGGCACATGCCATGTTGACC	4386

Right-hand alignment block (coordinate ranges approximately 4387–5714):

Label	Start	...	End
Qy	4575	TTGACCCTACTTTCACCATGGTGTTCGTGTGTGTGCGGGGTTCAGCAATAGTTAAAGGCC	4634
Db	4387	TTGACCCTACTTTCACCATGGTGTTCGTGTGTGTGCGGGGTTTCAGCAATAGTTAAAGGCC	4446
Qy	4635	AGCGTAGGGGGCCGCACAGGCCGTGGGAGAGCTGGCATATACTACTATGGGGAGTT	4694
Db	4447	AGCGTAGGGGGCCGCACAGGCCGTGGGAGAGCTGGCATATACTACTATGTGACGGGAGTT	4506
Qy	4695	GTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCGTTGACGCAGCGCCAAGG	4754
Db	4507	GTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTTGACGCAGCGCCAAGG	4566
Qy	4755	CATGGTATGGTTGTCATCAACAGAGCTCAAACATTCTGACACCTATCCGACACCCACCCAAAC	4814
Db	4567	CATGGTATGGTTGTCATCAACAGAGCTCAAACATTCTGACACCTATCCGACACCCACCCAAAC	4626
Qy	4815	CTGGGTTACCTGCGATAGGAGGCAAATTGGACGAGTGGGCTGATCTCTTTCTCATGGTCA	4874
Db	4627	CTGGGTTACCTGCGATAGGAGGCAAATTGGACGAGTGGGCTGATCTCTTTCTCATGGTCA	4686
Qy	4875	ACCCCGAACCTTCATTGTCAATACTGCAAAAAGAACTGCTGACAATTATGTTTTGTTGA	4934
Db	4687	ACCCCGAACCTTCATTGTCAATACTGCAAAAAGAACTGCTGACAATTATGTTTTGTTGA	4746
Qy	4935	CTGCAGCCCAACTACAACTGTGTCATCAGTATGGCATGCGTCTCCCAATGACGCACCAC	4994
Db	4747	CTGCAGCCCAACTACAACTGTGTCATCAGTATGGCATGCGTCTCCCAATGACGCACCAC	4806
Qy	4995	GGTGGGCAGGGGAGCCCGGCTTGGGGAAAAAACCTTCTGTGCGGTCGTTGGTGCGGCG	5054
Db	4807	GGTGGGCAGGGGAGCCCGGCTTGGGGAAAAAACCTTCTGTGCGGTCGTTGGTGCGGCT	4866
Qy	5055	CTGACGGCTGTCCTGCCCCAGAGCCCAGCGAGTGACGGAGATACCAAAATGTGTTCACTG	5114
Db	4867	GTGACGGCTGTCCTGCCCCAGAGCCCAGCGAGTGACGGAGATACCAAATGTGCTTCACTG	4926
Qy	5115	AAGTCAATAACTTCTGGGACACAGCCGCCACTCGCTGTGGGCTATGGCTATGTGGCTATATC	5174
Db	4927	AAGTCAATAACTTCTGGGACACAGCCGCCACTCGCTGTGGGCTATGGCTATGTGGCTATATC	4986
Qy	5175	TAGCCATTGACACTTTTGGGCGGCCGTTGTGTGCGGCGTTGTGTCTATTACATCAGTCC	5234
Db	4987	TAGCCATTGACACTTTTGGGCGGCCGTTGTGTGCGGCGTTGTGTCTATTACATCAGTCC	5046
Qy	5235	CTACCGGTGCTACTGTCGCGCCCCAGTGGTTGACGAAGAGAAATCGTGAGGAGTGTGCAT	5294
Db	5047	CTACCGGTGCTACTGTCGCGCCCCAGTGGTTGACGAAGAGGAGCAAATCGTGAGGAGTGTGCAT	5106
Qy	5295	CATTCATTCCCTTGGAGGCCATGCGTGTTTCATTGCGCCGTATTACTACCCCACTCCACCA	5354
Db	5107	CATTCATTCCCTTGGAGGCCATGCGTGTTTCATTGCGCCGTATTACTACCCCACTCCACCA	5166
Qy	5355	CAACTAGTCCTTCCATTCACATTGGAAACCGCCCCTGAAAAACTAACACCTTCTTGGGCCTC	5414
Db	5167	CAACTAGTCCTTCCATTCACATTGGAAACCGCCCCTGAAAAACTAACACCTTCTTGGGCCTC	5226
Qy	5415	ATGCAGCTACAATCCTGCTATCATCATAGAGTATTGCTGTGGTTTAGTCACTTTAGTCACTTACCTGACA	5474
Db	5227	ATGCAGCTACAATCCTGCTATCATCATAGAGTATTGCTGTGGTTTAGTCACTTTAGTCACTTACCTGACA	5286
Qy	5475	ATCCCTTTGCATCATGCGTGTTTGCTTTCATTGCGCGGTTATTACTACCCCACTACCTCACA	5534
Db	5287	ATCCCTTTGCATCATGCGTGTTTGCTTTCATTGCGCGGTTATTACTACCCCACTACCTCACA	5346
Qy	5535	AGATCAAAATGTTCCTGTCATTCATTGGGAGGGGCAATTGCGTCCAAGCTTACAGACGCTA	5594
Db	5347	AGATCAAAATGTTCCTGTCATTCATTGGGAGGGCAATTGCGTCCAAGCTTACAGACGCTA	5406
Qy	5595	GAGGGCACTGCGGTTGCGATGGCGCGGTCGCGGGGGCTGCGGAACAGCTCTTGGTACATGGACAT	5654
Db	5407	GAGRCGGACTGCGGTTGCGATGGCGCGGTCGCGGGGGCTGCGGAACAGCTCTTGGTACATGGACAT	5466
Qy	5655	CCGTGGGTTTTGTCTTTGACATGCTAGGCGGCTATGCTGCCGCCTCATCCACTGCTTGCT	5714

```
Qy    7872  GTCTGTTGTCGATGTGGTCACCCGATGCAGTCGGAGCCACATGGGATACAGTGTGTTTTG  7931
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7687  GTCTGTTGTCGATGTGGTCACCCGATGCAGTCGGAGCCACATGGGATACAGTGTGTTTTG  7746

Qy    7932  ACAGTACCATCACACCCGAGGATATCATGGTGAGACAGACATCTCACTCAGCAGCTAAAC  7991
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7747  ACAGTACCATCACACCCGAGGATATCATGGTGAGACAGACATCTCACTCAGCAGCTAAAC  7806

Qy    7992  TCAGTGACCAACACCAGGCTGGCATTCACACCATTGCGAGGCAGTTATACGCTGGAGGAC  8051
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7807  TCAGTGACCAACACCAGGCTGGCATTCACACCATTGCGAGGCAGTATCACGCTGGAGGAC  7866

Qy    8052  CGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTAGGTGTCTTCCGGCGTCT  8111
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7867  CGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTAGGTGTCTTCCGGCGTCT  7926

Qy    8112  ATACTACCTCAAGTTCAACAGTTTGACCTGCTGGCTGAAAGTAAATGCTGCAGCCGAAC  8171
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7927  ATACTACCTCAAGTTCAACAGTTTGACCTGCTGGCTGAAAGTAAATGCTGCAGCCGAAC  7986

Qy    8172  AGGCTGGCATGAAGAACCCTCGCTTCCTTATTTGCGGCGATGATTGCACCGTAATTTGGA  8231
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7987  AGGCTGGCATGAAGAACCCTCGCTTCCTTATTTGCGGCGATGATTGCACCGTAATTTGGA  8046

Qy    8232  AGAGCGCCGGAGCGCAGATGCAGACAAACAAGCAATGCGTGTCTTGCTGAGTGGATGAAGG  8291
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8047  AGAGCGCCGGAGCGCAGATGCAGACAAACAAGCAATGCGTGTCTTGCTGAGTGGATGAAGG  8106

Qy    8292  TGATGGGTGCACCACAAGATGTGTGCCTCAACCCCAAATACAGTTTGGAAGAATTAACAT  8351
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8107  TGATGGGTGCACCACAAGATGTGTGCCTCAACCCCAAATACAGTTTGGAAGAATTAACAT  8166

Qy    8352  CATGCTCATCAAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTACTACTTTCTTA  8411
            ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db    8167  CATGCTCATCAAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTACTACTTTCTTA  8226

Qy    8412  CAAGAGATCCTCGTATCCCCCTTGGCCAGGTGCTCGCCGAGGTCTGGGATACAACCCCA  8471
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8227  CAAGAGATCCTCGTATCCCCCTTGGCCAGGTGCTCGCCGAGGTCTGGGATACAACCCCA  8286

Qy    8472  GTGCTCGCGTGATTGGGTATCTAATACATCACTACCCCATGTTTGTGGCCGTGTGT  8531
            |:|||||||||||||||||||||||||||||||||| :||||||||||||||||||
Db    8287  GKGCKGCGTGATTGGGTATCTAATACATCACTACCCCATGTTTGTGGCCGTGTGT  8346

Qy    8532  TGGGCTGTCCATTTCATGGAGCAGAGATGCTCTTTGAGGACAAAACTTCCCGAGACTGTGACCT  8591
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8347  TGGGCTGTCCATTTCATGGAGCAGAGATGCTCTTTGAGGACAAAACTTCCCGAGACTGTGACCT  8406

Qy    8592  TTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAGCATCATTGCTG  8651
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8407  TTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAGCATCATTGCTG  8466

Qy    8652  GTGTGCACGGTATTGAGGGCTTTCTCCGGTGTGCGCGTACACCAACGCTGAGACTCCTCAGAG  8711
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8467  GTGTGCACGGTATTGAGGGCTTTCTCCGGTGTGCGCGTACACCAACGCTGAGATCCTCAGAG  8526

Qy    8712  TTTCCCAATCACTAACAGACATGACCAATGCCCCCCTGCGAGCCTTGGCGGGAAAGAAAGCCA  8771
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8527  TTTCCCAATCACTAACAGACACATGACCAATGCCCCCCTGCGAGCCTTGGCGGGAAAGAAAGCCA  8586

Qy    8772  GGGCGGTCCTCGCCAGCGCCAAGAGGCGTGGCGGAGCACACGCAAAATTGGCTCGCTTCC  8831
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8587  GGGCGGTCCTCGCCAGCGCCAAGAGGCGTGGCGGAGCACACG-AAAATTGGCTCGCTTCC  8645

Qy    8832  TTCTCTGGCATGCTACATCTACAGACCTTCACCAGGATAAGACGAGCGTGGCTCGGT  8891
            ||||||||| ||||||||||||||||| |||||||||||||||||||||||||||||
Db    8646  TTCTCTGGCATGCTACATCTACAGACCTTCACCAGGATTAGACGAGCCGTGGCTCGGT  8705

Qy    8892  ACACCACTTTCAATTATTGTGATGTTTACTCCCCCGGAGGGGATGTGTTTATTACCACAC  8951
            |||||||||||||||||||||||||||||| ::|||||| ||||||||||||||||||||
Db    8706  ACACCACTTTCAATTATTGTGATGTTTACT-CCCSGAGRGGGATGTGTTTATTACCACAC  8764
```

```
Qy    8952  AGAGAAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTGTCATTGTTTTGCCCTAGGGC  9011
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8765  AGAGAAGATTGCAGAAGTTTCTTGTGAAGTATTTGGCTGTGTCATTGTTTTGTGCCCTAGGGC  8824

Qy    9012  TCATTGCTGTTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAACTAACAGTTTT  9071
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8825  TCATTGCTGTTGGACTAGCCATCAGCTGAACCCCCAAATTCAAAATTAATTAACAGTTTT  8884

Qy    9072  TTTTTTTTTTTTTTT  9090
            |||||||||||||||
Db    8885  TTTTTTTTTTTTTTT  8903
```

RESULT 11
US-08-488-446-11
; Sequence 11, Application US/08488446
; Patent No. 6558898
; GENERAL INFORMATION:
;   APPLICANT: JOHN N. SIMONS
;   APPLICANT: TAMI J. PILOT-MATIAS
;   APPLICANT: GEORGE J. DAWSON
;   APPLICANT: GEORGE G. SCHLAUDER
;   APPLICANT: SURESH M. DESAI
;   APPLICANT: THOMAS P. LEARY
;   APPLICANT: ANTHONY SCOTT MUERHOFF
;   APPLICANT: JAMES C. ERKER
;   APPLICANT: SHERI L. BUIJK
;   APPLICANT: ISA K. MUSHAHWAR
;   TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
;   TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
;   NUMBER OF SEQUENCES: 716
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: ABBOTT LABORATORIES   D377/AP6D
;   STREET: 100 ABBOTT PARK ROAD
;   CITY: ABBOTT PARK
;   STATE: IL
;   COUNTRY: USA
;   ZIP: 60064-3500
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/488,446
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/08/424,550
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: POREMBSKI, PRISCILLA E.
;   REGISTRATION NUMBER: 33,207
;   REFERENCE/DOCKET NUMBER: 5527.PC.01
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 708-937-6365
;   TELEFAX: 708-938-2623
;   INFORMATION FOR SEQ ID NO: 11:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 8912 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-08-488-446-11

Query Match   92.5%; Score 8692.6; DB 4; Length 8912;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 8781; Conservative 56; Mismatches 52; Indels 10; Gaps 8;

```
Qy    196  TGGGTTCGGTGGTGGTGGCGCTTTAGGCAGCCTTCCACGCCCACCACCTCCCAGATAGAGC  255
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11  TGGGTTCGGTGGTGGTGGCGCTTTAGGCAGCCTTCCACGCCCACCACCTCCCAGATAGAGC  70
```

us-10-009-002-1.rni

Right column:

```
Qy  1336 CTTGACCAAACTGGCTTCACAAGTACCATATACGGCTATTGCGACTATGTTTAGCAGTGTACA 1395
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1151 CTTGACCAAACTGGCTTCACAAGTACCATATACGGCTATTGCGACTATGTTTAGCAGTGTACA 1210

Qy  1396 CTACCTGGCGGTTGGCGTCTGATCTCTGTACTACTAYGCCTCTCGGGGCAAGTGGTATCAGTTGCT 1455
         ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db  1211 CTACCTGGCGGTTGGCGTCTGATCTCACTACTAYGCCTCTCGGGGCAAGTGGTATCAGTTGCT 1270

Qy  1456 CCTAGCGCTTATGCTTACATAGAAGCGACCTCTGGAAACCCATCAGGGTGCCCACTGG 1515
         ||||||||||||||||||||  : ||||||||||||||||||||||||||||||||||
Db  1271 CCTAGCGCTTAYGCTTACATAGAACCCYATCAGGGTGCCCACTGG 1330

Qy  1516 ATGCTCAATAGCTGAGTTTGCTCGCCTTTGATGATACCATGTCCTTGCCACTCTTATTT 1575
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1331 ATGCTCAATAGCTGAGTTTGCTCGCCTTTGATGATACCATGTCCTTGCCACTCTTATTT 1390

Qy  1576 GAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGACCAGGCCATCACTCT 1635
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1391 GAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGACCAGGCCTGTCACTCT 1450

Qy  1636 AGAGTATATACAACTCCATATCTTGGTACCCCTATACAATCCCTGGTGCGAGAGGGATGTAT 1695
         |||||||||||: :|||||||||||||||||||||||||||||||||||||||||||||||
Db  1451 AGAGTATATABAACTCCATATCTTGGTACCCCTATACAATCCCTGGTGCGAGAGGGATGTAT 1510

Qy  1696 GGTTAAATTCAAAAATAACACATGGGTTGCTGCGTATTCGCAATGTGCCATCGTACTG 1755
         |||||||||||||||||||||||||||||||||| : |||||||||||||||||||||
Db  1511 GGTTAAATTCAAAAATAACACATGGGTTGCTGCCG-WWTCGGCAATGTGCCATCGTACTG 1569

Qy  1756 CACTATGGGCACTGATGCAGTGTGGAAGCACACTCGCAACACTTACGAAGCATGCGGTGT 1815
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1570 CACTATGGGCACTGATGCAGTGTGGAASSACAGTCGCAACACTTACGAAGCATGCGGTGT 1629

Qy  1816 AACACCATGGCTAACAACCGCATGGCACAACCGCATGGCACAACCGCTCAGCCCTGAAATTGGCTATATTACA 1875
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1630 AACACCATGGCTAACAACCGCATGGCACAACCGCTCAGCCCTGAAATTGGCTATATTACA 1689

Qy  1876 ATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTCAGGCCCATTGTATTT 1935
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1690 ATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTCAGGCCCACTTGTATTT 1749

Qy  1936 TGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAATTCCACTCTTCCTACC 1995
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1750 TGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCTGTGAATTCCACTCTTCCTACC 1809

Qy  1996 ACCGGAGAGGTGGCGTAGGTGGCGTAGGTTTACAGTGATGTGAAAGACCTAGCCACAGGATTGATCACCAAAGA 2055
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1810 ACCGGAGAGGTGGCGTAGGTTTACAGTGATGTGAAAGACCTAGCCACAGGATTGATCACCAAAGA 1869

Qy  2056 GGTTCCGCAAGGGTTTACAGTGATGTGTGTTCTTTGTCTTTGTCCTTAGCAAGTATCTTTAT 2115
         ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1870 GGTTCCGCAA--GGTTTACAGTGATGTGTGTTCTTTGTCTTTGTCCTTAGCAAGTATCTTAT 1928

Qy  2116 CAAAGCCTGGAAAAAATTATCAGGTCTTATATTCCGCAGCCAAGTATCTTAT 2175
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1929 CAAAGCCTGGAAAAAATTATCAGTYCTTATATTCCGCAGCCAAGTATCTTAT 1988

Qy  2176 AGTTACCACCAAGGCCGTGGTGCTAATTCGTTGGGGGTTGTGTGGCAGCAAGTATCTTAT 2235
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1989 AGTTACCACCAAGGCCGTGGTGCTAATTCGTTGGGGGTTGTGTGGCAGCAAGTATCTTAT 2048

Qy  2236 TTTAGCCTTACCTCTGTTGTCCCCTTTGTTGTTTGGGGCGGCCTTCGTGGGGCCGCCTTACCCTTTGCG 2295
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2049 TTTAGCCTTACCTCTGTTGTCCCCTTTGTTGTTTGGGGCGGCCTTCTCGGTTACMCTTTGCG 2108

Qy  2296 TCCTGTGCTCCCATCCCAGTCGTATCTCCAAGTGGCTGGCTGGATGTTTTGTCTAAAGCTCA 2355
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2109 TCCTGTGCTCCCATCCCAGTCGTATCTCCAAGTGGCTGGCTGGATGTTTTGTCTAAAGCTCA 2168

Qy  2356 AGTAGCTCCTTTGCTTTGATTTTCTTCATCTGTGCTATCTCCGCTGCAGGCTACGTTA 2415
         |||||||||||: ||||||||||||||||||||||||||||||||||||||||||||||
Db  2169 AGTAGCTCMTTTGCTTTGATTTTCTTCATCTGTGCTATCTCCGCTGCAGGCTACGTTA 2228

Qy  2416 TGCTGCCCTTTTAGGGTTTGTGCCCATGCGTGCGGGGCTTGCCCCTAACTTTCTTTGTTGC 2475
         ||||||
Db                                                                 
```

Left column:

```
Qy  256  GGCGGCACTGTGAGGGAAGACCCGGGGACCGGTCACTACCAAGGACGCAGACCTCTTTTTGA 315
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  71   GGCGGCACTGTGAGGGAAGACCCGGGGACCGGTCACTACCAAGGACGCAGACCTCTTTTTGA 130

Qy  316  GTATCACGCCTCCGGAGTAGTTGGGCAAGCCCACCTATATGTGTTGGGATGGTTGGGGT 375
         ||||||||||||||||||||||||||||||||||||| : ||||||||||||||||||||
Db  131  GTATCACGCCTCCGGAGTAGTTGGGCAAGCCCACCTAYCACTAYGTGTTGGGATGGTTGGGGT 190

Qy  376  TAGCCATCCATACCGTACTGGCCTGATAGGGTCCTTGCCGGAGTCTGGGAGTCTCGTAG 435
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  191  TAGCCATCCATACCGTACTGGCCTGATAGGGTCCTTGCCGGAGTCTGGGAGTCTCGTAG 250

Qy  436  ACCGTAGCACATGCCTGTGTTATTTCTACTCAAACAAGTCCTGTACCTGCGCCCAGAACGCG 495
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  251  ACCGTAGCACATGCCTGTGTTATTTCTACTCAAACAAGTCCTGTACCTGCGCCCAGAAACGCG 310

Qy  496  CAACAACAAGCAGACGGCAAAGCGCAGGCGCTTCATATCTGTGTCCATTAAAACATCTGTTGAAAGGGG 555
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  311  CAACAACAAGCAGACGGCAAAGCGCAGGCGCTTCATATCTGTGTCCATTAAAACATCTGTTGAAAGGGG 370

Qy  556  ACAACGAGCAAAGTCCAGCGCGATGCTCGGCCGCTCGGCTGCTTTGCCAGCTCATGGTTGGGGACGG 615
         ||||||||||||||||||| : ||| : |||||||||||||||||||||||||||||||||
Db  371  ACAACGAGCAARGCGCAAAGTCCAGCGCGATGCTCGGCCGCTCGGCTGCTTTGCCAGCTCATGGTTGGGGACGG 430

Qy  616  TATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTGCCAGTCATGGTTGGGGACGG 675
         |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db  431  TATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTRCCAGTCATGGTTGGGGACGG 490

Qy  676  CCAAGACCCCTGCCCATAAGTCTCGCAATCTTCTGGAATCTTCTGGATTACCCCTTGGGGTG 735
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  491  CCAAGACCCCTGCCCATAAGTCTCGCAATCTTCTGGAATCTTCTGGATTACCCCTTGGGGTG 550

Qy  736  GATTGGTGATGTTACAACTCACACACCTCTCAGTAGGGCCCGCTGGTGGCAGGAGCGGTCGT 795
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  551  GATTGGTGATGTTACAACTCACACACCTCCAGTAGGGCCCGCTGGTGGCAGGAGCGGTCGT 610

Qy  796  TCGACCAGTCTGCCAGATAGTACGCTTGCTGGAGGATGGAGTCAACTGGGCTACTGGTTG 855
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  611  TCGACCAGTCTGCCAGATAGTACGCTTGCTGGAGGATGGAGTCAACTGGGCTACTGGTTG 670

Qy  856  GTTTATCTTATTGTCTCCTTCCCACTTGCCTACGTCGTATGTCGTCGTATGTCTGCCAATTCCTGACGAGCCTAGTGTGCCGGA 915
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  671  GTTTATCTTATYGTCTCCTTCCCACTTGCCTACGTCGTATGTCGTCGTATGTCTGCCAATTCCTGACGAGCCTAGTGTGCCGGA 730

Qy  916  GCGGGTCACTGACCCTGGGTTCCCGGCCAATCGTACACGACACAAATACCACAATCCCAATTGCTGCCCAATTCCTGACGAGCCGTAATCA 975
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  731  GCGGGTCACTGACCCGGGTTCCCGGCCAATCGTACACGACACAAATACCACAATCCTCCAATTGCTGCCGTAATCA 790

Qy  976  GGTTATCTATTGTTCTCCTTCCTTCCCACTTGCCTACACGTTGTTGTGTGGATCTGTGCCGGA 1035
         |||||||  ||||||||||||||||||||||||||||||||||||||||| : ||||||||
Db  791  GGTTATCTAYTGTTCTCCTTCCTTCCCACTTGCCTACACGTTGTTGTGTGTGYGGA 850

Qy  1036 CGAGTGCTGGGGTTCCCGGCCAATCGTACGCCAATCGTACACACACCTTCACACCGGTCGGGTGGCACGGA 1095
         |||||||||||||||||||||||||||||||: : |||||: |||||||||||||||||
Db  851  CGAGTGCTGGGGTTCCCGGCCAATCGTACGCCGRTACACATCCGRTACACACACCTTCACACCGGTCGGGTGGCACGGA 910

Qy  1096 CTCCTTCTTGGCTGACCAATTGATTTTTGTTATGGGCGCTCTGTTATGGGCGCTGTGACGCGCCCT 1155
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  911  CTCCTTCTTGGCTGACCAATTGATTTTTGTTATGGGCGCTCTGTTATGGGCGCTGTGACGCGCCCT 970

Qy  1156 TGACATTGGTAGGTTGTGTGGGTCGTGGTCGTGTATTAGTAGTCCGTTGTCCTTGTCGGTCAGGGCACTG 1215
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  971  TGACATTGGTAGGTTGTGTGGGTCGTGGTCGTGTATTAGTAGTCCGTTGTCCTTGTCGGTCAGGGCACTG 1030

Qy  1216 GCTTATTCACATAGACCTCAATGAAAACTGGTACTTGTTACCTGGAAGTGCCCACTGGAAT 1275
         |||||||||||||||||| : |||||||||||||||||||||||||||||||||||||||
Db  1031 GCTTATTCACATAGACCTCAATGAAAKTGCCTACTTGTTACCTGGAAKTGCCTACTGGAAT 1090

Qy  1276 AGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGCGCCAAGGCCGCGCCAAGGCTGTCATCTT 1335
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1091 AGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGCGCCAAGGCCGCGCCAAGGCTGTCATCTT 1150
```

Qy

Db

This page consists of a two-part nucleotide sequence alignment (query "Qy" versus database "Db") with match bars between the rows and coordinate numbers at the ends of each line. The readable coordinate markers and strand labels are listed below.

Right-hand block (Db / Qy lines with start and end coordinates):

Label	Start	End
Db	3307	3366
Qy	3555	3614
Db	3367	3426
Qy	3615	3674
Db	3427	3486
Qy	3675	3734
Db	3487	3546
Qy	3735	3794
Db	3547	3606
Qy	3795	3854
Db	3607	3666
Qy	3855	3914
Db	3667	3726
Qy	3915	3974
Db	3727	3786
Qy	3975	4034
Db	3787	3846
Qy	4035	4094
Db	3847	3906
Qy	4095	4154
Db	3907	3966
Qy	4155	4214
Db	3967	4026
Qy	4215	4274
Db	4027	4086
Qy	4275	4334
Db	4087	4146
Qy	4335	4394
Db	4147	4206
Qy	4395	4454
Db	4207	4266
Qy	4455	4514
Db	4267	4326
Qy	4515	4574
Db	4327	4386
Qy	4575	4634
Db	4387	4446

Left-hand block (Db / Qy lines with start and end coordinates):

Label	Start	End
Db	—	2288
Qy	2476	2535
Db	2289	2348
Qy	2536	2595
Db	2349	2407
Qy	2596	2654
Db	2408	2466
Qy	2655	2714
Db	2467	2526
Qy	2715	2774
Db	2527	2586
Qy	2775	2834
Db	2587	2646
Qy	2835	2894
Db	2647	2706
Qy	2895	2954
Db	2707	2766
Qy	2955	3014
Db	2767	2826
Qy	3015	3074
Db	2827	2886
Qy	3075	3134
Db	2887	2946
Qy	3135	3194
Db	2947	3006
Qy	3195	3254
Db	3007	3066
Qy	3255	3314
Db	3067	3126
Qy	3315	3374
Db	3127	3186
Qy	3375	3434
Db	3187	3246
Qy	3435	3494
Db	3247	3306
Qy	3495	3554

```
Left column

Qy  4635 AGCGTAGGGGCCGCACAGGCCGTGGGAGAGCTTGGCATATACTACTATGTGAGACGGAGTT 4694
               ||||||||| |||||||||||||||||||| ||||||||||||||||||| ||||||||||
Db  4447 AGCGTAGGGGCCCACAGGCCGTGGGAGAGCTTGGCATATACTACTATGTGAGACGGAGTT 4506

Qy  4695 GTACCCCCTTCGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTTCGACGCAGCCAAGG 4754
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4507 GTACCCCTTCGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTTCGACGCAGCCAAGG 4566

Qy  4755 CATGGTATGGTTGTCATCAACAGAAGCTCAAACTATTCTGGACACTTATCGCCACCCAAC 4814
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4567 CATGGTATGGTTGTCATCAACAGAAGCTCAAACTATTCTGGACACTTATCGCCACCCAAC 4626

Qy  4815 CTGGGTTACCTGCGATAGGAGCAAATTGGACGAGTGGGCTGATCTCTTTCTATGGTCA 4874
               ||||||||||||||||||||| |||||| |||||||||||||||||||||||||||||
Db  4627 CTGGGTTACCTGCGATAGGAGCAAATTGGACGAGTGGGCTGATCTCTTTCTATGGTCA 4686

Qy  4875 ACCCGAACCTTCATTTGTCAATACTGTGCAAAAGAACTGCTGACAATTATGTTTGTTGA 4934
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4687 ACCCGAACCTTCATTTGTCAATACTGTGCAAAAGAACTGCTGACAATTATGTTTGTTGA 4746

Qy  4935 CTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCCCAATGACGCACCAC 4994
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4747 CTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCCCAATGACGCACCAC 4806

Qy  4995 GGTGGCAGGGGAGCCCGCTTGGGAAAAACCTGTGTGGGGTTCGTGGCGCTTGGACGGCG 5054
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4807 GGTGGCAGGGGAGCCCGCTTGGGAAAAACCTGTGTGGGGTTCGTGGCGCTTGGACGGCT 4866

Qy  5055 CTGACGCGCCGTGCTCGGCCCAGAGCCCAGCGGAGTGACGAGTGACGTACCAAATGTCACTG 5114
               |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4867 GTGACGCGCCGTGCTCGGCCCAGAGCCCAGCGGAGTGACGAGTGACGTACCAAATGTCACTG 4926

Qy  5115 AAGTCAATACTTCTGGAGGCCATGGTTGCTGTCATCGCTTGGCGTTGGAGTGGCTATGGCTTATC 5174
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4927 AAGTCAATACTTCTGGAGGCCATGGTTGCTGTCATCGCTTGGCGTTGGAGTGGCTATGGCTTATC 4986

Qy  5175 TAGCCATTGACACTTTTGGCGCCACTTGTGTGCGCGCGTTGCTTATTACATCAGTCC 5234
               |||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  4987 TAGCCATTGACACTTTTGGCGCCACTTGTGTGCGCGCGTTGCTTATTACATCAGTCC 5046

Qy  5235 CTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAGAAGAAATCGTTGGAGGAGTGTGCAT 5294
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5047 CTACCGGTGCTACTGTCGCCCCAGTGGTTGACGAGAAGAAATCGTTGGAGGAGTGTGCAT 5106

Qy  5295 CATTCATTCCCTTGGAGGCCATGGCCATGTTGCTTTCATTGCGGGTATTACCCCACTTCACCA 5354
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5107 CATTCATTCCCTTGGAGGCCATGGCCATGTTGCTTTCATTGCGGGTATTACCCCACTCACCA 5166

Qy  5355 CAACTAGTCCTTTCACATTGAAAACCGCCCCTTGAAAAACTTAACACCTTTCTTGGGCCTC 5414
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5167 CAACTAGTCCTTTCACATTGAAAACCGCCCCTTGAAAAACTTAACACCTTTCTTGGGCCTC 5226

Qy  5415 ATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGGTGGTTTAGTCACTTTACCCTGACA 5474
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5227 ATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGGTGGTTTAGTCACTTTACCCTGACA 5286

Qy  5475 ATCCCTTTGCATCATGCATGATGGCCCGGGCCGCTGCGGAACTCGTATTACTACCCCACTCACA 5534
               |||||| ||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db  5287 ATCCCTTTGCATCATGCATGATGGCCCGGGCCGCTGCGGAACTCGTATTACTACCCCACTCACA 5346

Qy  5535 AGATCAAAATGTTCCTGTCATTATTTGGAGGGGCAATTGCGTCAACTTAACACCTTTACAGACGCTA 5594
               |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5347 AGATCAAAATGTTCCTGTCATTATTTGGAGGGGCAATTGCGTCAACTTAACACCTTTACAGACGCTA 5406

Qy  5595 GAGGCGCACTGGCGTTTGTTCTTGCTTGGACTGGGGCCCTGCGGAACAGAGCTCTTGGAGGGCTTACCAGACGCTTGGACGGTCTCCGGCTCTTGTTT 5654
               |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5407 GAGGCGCACTGGCGTTTGTTCTTGCTTGGACTGGGGCCCTGCGGAACAGAGCTCTTGGACTGTGTCYGGGCGTGTCCGGCTCTTGTTT 5466

Qy  5655 CGGTGGGTTTGTCTTGACATGTGACATGGCGCTAGGCGCTATGCGCCGCCTTCATCCATGCTGCT 5714
               |||||||||||||||||||||||| ||||| |||||||||||||||||||||||||||||
Db  5467 CGGTGGGTTTGTCTTGACATGTGACATGGCGCTAGGCGCATGCGCCGCCTTCATCCATGCTGCT 5526
```

```
Right column

Qy  5715 TGACATTTAAATGCTTGATGGTGGTGCCCACTATGGATCAGCTTGCTGGTTTAGTCT 5774
               |||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  5527 TGACATTTAAATGCTTGATGGTGGTGAGTGGCYCACTATGGATCAGCTTGCTGGTTTAGTCT 5586

Qy  5775 ACTCCGCGTTCAATCCGCCGCAGGAGTTGTGGCGTCTTGTGTCAGCTTGTGTCAATGTTTG 5834
               |||||||||||||||||||||||||||||||||||||||||||||||:||||||| |||||
Db  5587 ACTCCGCGTTCAATCCGCCGCAGGAGTTGTGGCGTCTTGTGTCAGCTTGTGTCAATGTTTG 5646

Qy  5835 CTTTGACAACAGCAGGGCCAGATCACTGCGCCCAACAGACTTCTTACTACTGCTTGCTAGGA 5894
               |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db  5647 CTTTGACAACAGCAGGGCCAGATCACTGCGCCCAACAGACTTCTTACTACTGCTTGCTAGGA 5706

Qy  5895 GCAACACTGTATGTAATGACTACTTATTGCCACTCGTGACATCCGCAGGAAGADATACTGG 5954
               |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db  5707 GCAACACTGTATGTARTGACTACTTATTGCCACTCGTGACATCCGCAGGAAGADATACTGG 5766

Qy  5955 GCATTCTGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCATCCGTTGGCTCCACA 6014
               ||||||||||||||||||||||||||||||||||||:|||| ||||||||||||||CACA
Db  5767 GCATTCTGGAGGCATCTACCCCCTGGAGTRTCATATCAGCTTGCATCCGTTGGCTYCACA 5826

Qy  6015 CCCCGACGGAGGATGATTGCGGCCTCAGTTGCTTGGGGTCTAGAGATTTGGCAGTATGTGT 6074
               |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db  5827 CCCCGACGGAGGATGATTGCGGCCTCAGTTGCTTGGGGTCTARAGATTTGGCAGTATGTGT 5886

Qy  6075 GCAATTCTTGTGTGATTTGCTTGTCCTTAAAGTGGAGTTCAGAGCATGTTAACA 6134
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5887 GCAATTCTTGTGTGATTTGCTTGTCCTTAAAGTGGAGTTCAGAGCATGTTAACA 5946

Qy  6135 TTCCTGGTTGTCTTTCACAGCTGCCAGAGGGGTACAAGGGCCCTGGATTGGATCAG 6194
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5947 TTCCTGGTTGTCTTTCACAGCTGCCAGAGGGGTACAAGGGCCCTGGATTGGATCAG 6006

Qy  6195 GTATGCTCCAAGCACGCTGTCCATGCGGTGCTGAACTCATCTTTCTGGAGAATGGTT 6254
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6007 GTATGCTCCAAGCACGCTGTCCATGCGGTGCTGAACTCATCTTTCTGGAGAATGGTT 6066

Qy  6255 TTGCAAAACTTTACAAAGGACCCAGAACTGTTCAAATTACTGGAGAGGGCGTGTTCCAG 6314
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6067 TTGCAAAACTTTACAAAGGACCCAGAACTGTTCAAATTACTGGAGAGGGCGTGTTCCAG 6126

Qy  6315 TCAACGCTAGGCTGTGTGTGGTGGTGGCTAGACCGATTGGACTAGTCTTGTCG 6374
               ||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db  6127 TCAACGCTAGGCTGTGTGTGGTGGTGGCTAGACCGATTGGACTAGTCTTGTCG 6186

Qy  6375 TCAATTATGGCGTTAGGGACTACTGTAAATATGTAAATATGGGAGAAAATGGGAGATCACATTTTTTGTTA 6434
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6187 TCAATTATGGCGTTAGGGACTACTGTAAATATGTAAATATGGGAGAAAATGGGAGATCACATTTTTTGTTA 6246

Qy  6435 CAGCAGTATCCTCTCCAAATGTCGTTCACCCAGGTGCGCCCAACCTTGAGAGCTGCAG 6494
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6247 CAGCAGTATCCTCTCCAAATGTCGTTCACCCAGGTGCGCCCAACCTTGAGAGCTGCAG 6306

Qy  6495 TGGCCGTGGACGGCGTACAGGTTCAGTGTAATCTCAGGGTAAGCGGTAAAACTCCTTGGACGA 6554
               ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  6307 TGGCCGTGGACCGGCGTACAGGTTCAGYGTTATCTCAGGGTAAGCGGTAAAACTCCTTGGACGA 6366

Qy  6555 CATCTGCTTGCTGTCGTTACGGTGCTGTAAGCGGTAAAACTGTTAAGCTTCCCTTCCGCG 6614
               |||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db  6367 CATCTGCTTGCTGTCGTTACGGTGCTGTAAGCGGTAAAACTGTTAAGCTTCCCTTCCGCG 6426

Qy  6615 TTGACGGTCACACCTGTGTGCGCATGCAACTTAATTTGCGTGATGCACTTGAGACAA 6674
               |||||| ||||||||||||||| |||||||||||||||||||||||||||||||
Db  6427 TTGACGGACACACCTGTGTGCGTCATGCAACTTAATTTGCGTGATGCACTTGAGGCAA 6486

Qy  6675 ATGACTGTAATTCCACAACACTCCTAGTGATGAGGCCGCAGTGTCTCTTGTTT 6734
               ||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db  6487 ATGACTGTAATTTCCATAAACACACTCCCAGTGATGAGGCCGCAGTGTCTCTTGTTT 6546

Qy  6735 TCAAACAGGAGTTGCGGGGTACAAAACCAATTGCTTGAGGCAATTTCAGCTGGCGTTGACA 6794
               ||||||||||||||||||| ||||||||||:|||||||||||||||||||||||||||||
Db  6547 TCAAACAGGAGTTGCGGGGTACAAAACCAATTGCTTGAGGCAATTTCAGCTGGCGTTGACA 6606

Qy  6795 CCACCAAACTGCCAGCCCCCTCC--ATCGAAGAGGTAGTGGTAAGAAAGCGCCAGTCC 6851
               ||||||||||||||||||||||||  |||||||||||||||||||||||||||||||
Db  6607
```

Left column:

```
Db        ||||||||||||||||||||||
     6607 CCACCCAAACTGCCAGCCCCCTCCCAGATCGAAGAGGTAGTGGTAAGAAAGCGCCCAGTTCC 6666
Qy   6852 GGGCAAGAACTGGTTCGCTTACCTTGCCTCCTCCCCCTCCGAGATCCGTCCCAGGAGTGTCAT 6911
Db   6667 GGGCAAGAACTGGTTCGCTTACCTTGCCTCCTCCCCCTCCGAGATCCGTCCCAGGAGTGTCAT 6726
Qy   6912 GTCCTGAAAGCCTGCAACGAAGTGACCCGTTAGAGCTTCAAACCTCCTTCAC 6971
Db   6727 GTCCTGAAAGCCTGCAACGAAGTGACCCGTTAGAGCTTCAAMCCTCCTTCCTTCAC 6786
Qy   6972 CACCTGTTCTACAGTTGGCCATGCCGATGCCCGTGTGGGACGGGTGAGTGTAACCCTT 7031
Db   6787 CACCTGTTCTRCAGTTGGCCATGCCGATGCCCCTGTTGGGAGCAGGTGAGTGTAACCCTT 6846
Qy   7032 TCACTGCAATTGGATGTGCAATGACCGGAACAGGCCGGAGCCCTGATGATTACCCAGTT 7091
Db   6847 TCACTGCAATTGGATGTGCAATGACCGGAACAARGYGGAGKCCCCWMAKRATTTACCCAGTT 6906
Qy   7092 ACCCTCCCAAAAAGGAGGTCTCGAATGGTCAGACGAAAGTTGGTCGACGGCTACAACCG 7151
Db   6907 ACCCTCCCAAAAAGGAGGTCTCGAATGGTCAGACGAAAGTTGGTCAACGACTACAACCG 6966
Qy   7152 CTTCCAGCTACGTTACTGGCCCCCCGTACCCTAAGATACGGGAAAGGATTTCCACTCAGT 7211
Db   6967 CTTCCAGCTACGTTACTGGCCCCCCGTACCCTAAGATACGGGGCAAGGATTTCCACTCAAT 7026
Qy   7212 CAGCCCCCGCCAAACGGCCTACAAAGGAAGTTGGGAAAGAGAGTGAGTTTTCGTGCAGCA 7271
Db   7027 CAGCCACCGCCAAACGCCTACAAAGGAAGTTGGGAAAGAGAGTGAGTTTTCGTGCAGCA 7086
Qy   7272 TGAGCTACACCTGGACCGATGATTAGCTTCAAAACTGCTTCGTTCTGTCTGCAA 7331
Db   7087 TGAGCTACACCTTGGACCGATGATTAGCTTCAAAACTGCTTCGTTCTGTCTGCAA 7146
Qy   7332 CTCGGGCCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGTGACTGAGCCGC 7391
Db   7147 CTCGGGCCATCACTAGTGGTTTCCTCAAACAAAGATCATTGGTGTATGTGACTGAGCCGC 7206
Qy   7392 GGGGATGCGGAGCTTAGAAAACAAAAGTCACTATTAATAGACAACCTCTGTTCCCCCCAT 7451
Db   7207 GGGGATGCGGAGCTTAGAAAACAAAAGTCACTATTAATAGACAACCTCTGTTCCCCCCAT 7266
Qy   7452 CATACCACACAAGCAAGTGAGATTGGCTAAGGAAAAGCTTCAAAAGTTGTCGGTGTCATGT 7511
Db   7267 CATACCACACAAGCAAGTGAGATTGGCTAAGGAAAAGCTTCAAAAGTTGTCGGTGTCATGT 7326
Qy   7512 GGGACTATGATGAAGTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTCCCACATACTG 7571
Db   7327 GGGACTATGATGAAGTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTCCCACATCACTG 7386
Qy   7572 GCCTTCGGGGACTGATGTTCGTTCTGGAGCAGCCCGCAAACTGCCAAACTTCCAAAGG 7631
Db   7387 GCCTTCGGGGACTGATGTTCGTTCTGGAGCGGCCCCCAAACTGCCAAACTTCCAAAGG 7446
Qy   7632 AGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATGCGGCAAAACTGTAGTTGATAGTTCGT 7691
Db   7447 AGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATGCGGCAAAACTGTAGTTGATAGTCCGT 7506
Qy   7692 ACCCCCACCTTGTGAAGACCCCCAGAAAACCAACAAGAAACCCCCAAGGCTTATCTCGT 7751
Db   7507 ACCCCCACCTTGAAAATGAGATGTGTGAAGACCCCCAGAAAACCAACAAGAAACCCCCAAGGCTTATCTCGT 7566
Qy   7752 AGGAGGTCTTCGTGAAGACAGTGTTTACGGTTCAGGTTGCTCCTCGACG 7811
Db   7567 AGGAGGTCTTCGTGAAGACAGTGTTTACGGTTCAGGTTGCTCCTCGACG 7626
Qy   7812 TAGTTAAAGCTGTCATGGAGATGCGTACGGGTTGTGTAGATCGGGTTACCCGTGTCAAGC 7871
Db   7627 TAGTTAAAGCTGTCATGGAGATGCGTACGGGTTTGTAGATCCACGTGTACCCGTGTCAAGC 7686
Qy   7872 GTCTGTTGTGTCGATGTGGTCACCCGATGCCACATGCGATCGGAGCCCAGCACAGTGTGTTTG 7931
Db        ||||||||||||||||
```

Right column:

```
Db   7687 GTCTGTTGTGTCGATGTGGTCACCCGATGCCACATGCGATCGGAGCCCAGCACAGTGTGTTTG 7746
Qy   7932 ACAGTACCACCATCACACCCGAGGATATCATGTGAGACAGACATCTACTCAGCAGCTAAAC 7991
Db   7747 ACAGTACCACCATCACACCCGAGGATATCATGTGAGACAGACATCTACTCAGCAGCTAAAC 7806
Qy   7992 TCAGTGACCAACGAGCTGGCATTGCAGAGCTGCGAGGCAGTTATACGGTGGAGGAC 8051
Db   7807 TCAGTGACCAACGAGCTGGCATTCACACCATTCACCACCATTCAGCGCTGGAGGAC 7866
Qy   8052 CGATGATCGGCTTATGATGGCCGAGATCGGATATCGTAGGTGTAGGTGTTCCGGGTCT 8111
Db   7867 CGATGATCGGCTTATGATGGCCGAGATCGGATATCGTAGGTGTTCCGGGTCT 7926
Qy   8112 ATACTACCTCAAGTTCCAACAGTTGACCTGGCTGAAGGTAAATGCTGACCGGTAATTTGA 8171
Db   7927 ATACTACCTCAAGTTCCAACAGTTGACCTGGCTGAAGGTAAATGCTGACCGGTAATTTGA 7986
Qy   8172 AGGCTGCATGAAGAACCCTCGCTTCCTTATTGCGGCGATGATTGCACCGGTAATTTGA 8231
Db   7987 AGGCTGCATGAAGAACCCTCGCTTCCTTATTGCGGCGATGATTGCACCGGTAATTTGA 8046
Qy   8232 AGAGCGCCGGAGCAGATGCAGAACAAGCAATGCGTCTTCGTTGCTAGCTGATGATGAACGC 8291
Db   8047 AGAGCGCCGGAGCAGATGCAGAACAAGCAATGCGTCTTCGTTGCTAGCTGATGATGAACGC 8106
Qy   8292 TGATGGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACAGTTTGAAGAATTAACAT 8351
Db   8107 TGATGGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACAGTTTGAAGAATTAACAT 8166
Qy   8352 CATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTACTACTTTCTTA 8411
Db   8167 CATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTACTACTTTCTTA 8226
Qy   8412 CAAGAGATCCTCGTATCCCCTTGGCAGGTGCTCTGCCGAGGTCTGGGATACAACCCCA 8471
Db   8227 CAAGAGATCCTCGTATCCCCTTGGCAGGTGCTCTGCCGAGGTCTGGGATACAACCCCA 8286
Qy   8472 GTGCTGCGTGGATTGGGTATCTAATAATAGACAACCTCTGTTGTTAGCCGTGTGT 8531
Db   8287 GKGCKGGGTGGATTGGGTATCTAATAATAGACAACCTCTGTTGTTAGCCGTGTGT 8346
Qy   8532 TGGCTGTCCATTTCATCGGAGCAGATGCTCTTTGAGGACAAACTTCCCGAGACTGTGACCT 8591
Db   8347 TGGCTGTCCATTTCATCGGAGCAGATGCTCTTTGAGGACAAACTTCCCGAGACTGTGACCT 8406
Qy   8592 TTGACTGGTATGGGAAAAATTATACGGTGCCCTGTAGAAGATCTGCCCAGCATCATTGCTG 8651
Db   8407 TTGACTGGTATGGGAAAAATTATACGGTGCCCTGTAGAAGATCTGCCCAGCATCATTGCTG 8466
Qy   8652 GTGTGACGGTATTGGAGCTTTCTCGGGTCGCTACCAACGCTGAGATCCTCAGAG 8711
Db   8467 GTGTGACGGTATTGGAGCTTTCTCGGGTCGCTACCAACGCTGAGATCCTCAGAG 8526
Qy   8712 TTTCCCAATCACTAACACAGACATGACCATGCCCCCCCTGCGAGCCTGGCGGCGAAAAGCCA 8771
Db   8527 TTTCCCAATCACTAACACATGACCATGCCCCCCCTGCGAGCCTGGCGGCGAAAAGCCA 8586
Qy   8772 GGGCGGTCCTCGCATGCTACATCTAGACCTTCGT 8831
Db   8587 GGGCGGTCCTCGCATGCTACATCTAGACCTTCGT 8645
Qy   8832 TTCTCTGCATGCTACATCTAGACCTTCTAGATAGACGAGCCGTGCTGCTTCC 8891
Db   8646 TTCTCTGGCATGCTACATCTAGACCTTCACCCAGATTTGGATAAGACGAGCCGTGCTGGCTTCGGT 8705
Qy   8892 ACACCACTTTCAATTATTGTGATGTTTATTACTCCCCGAGGGGATGTGTTTATTACACCAC 8951
Db   8706 ACACCACTTTTCAATTATTGTGATGTTTATTACT-CCCSGAGRGGATGTGTTTATTACACCAC 8764
Qy   8952 AGAGAAGATTGCAGAAGTTCTTGTGAAGTTGTGCCTAGGGC 9011
Db   8765 AGAGAAGATTGCAGAAGTTTCTTGTGAAGTATTTGGCTGTCATTGTTTCATTGTTGCCCTAGGGC 8824
```

```
Qy    9012 TCATTGCTGTGTTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAAATTAACTAACAGTTTT 9071
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8825 TCATTGCTGTGTTGGACTAGCCATCAGCTGAACCCCCAAATTCAAAATTCAAAAATTAATTAACAGTTTT 8884

Qy    9072 TTTTTTTTTTTTTTTTTT 9090
           ||||||||||
Db    8885 TTTTTTTTTTTTTTTTTT 8903
```

RESULT 12
US-08-467-344A-11
; Sequence 11, Application US/08467344A
; Patent No. 6586568
   GENERAL INFORMATION:
      APPLICANT: JOHN N. SIMONS
               TAMI J. PILOT-MATIAS
               GEORGE J. DAWSON
               GEORGE G. SCHLAUDER
               SURESH M. DESAI
               THOMAS P. LEARY
               ANTHONY SCOTT MUERHOFF
               JAMES C. ERKER
               SHERI L. BUIJK
               ISA K. MUSHAHWAR
      TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
               REAGENTS AND METHODS FOR THEIR USE
      NUMBER OF SEQUENCES: 716
      CORRESPONDENCE ADDRESS:
         ADDRESSEE: ABBOTT LABORATORIES   D377/AP6D
         STREET: 100 ABBOTT PARK ROAD
         CITY: ABBOTT PARK
         STATE: IL
         COUNTRY: USA
         ZIP: 60064-3500
      COMPUTER READABLE FORM:
         MEDIUM TYPE: Floppy disk
         COMPUTER: IBM PC compatible
         OPERATING SYSTEM: PC-DOS/MS-DOS
         SOFTWARE: Patentin Release #1.0, Version #1.25
      CURRENT APPLICATION DATA:
         APPLICATION NUMBER: US/08/467,344A
         FILING DATE: 07-Jun-1995
         CLASSIFICATION: <Unknown>
      PRIOR APPLICATION DATA:
         APPLICATION NUMBER: 08/424,550
         FILING DATE: <Unknown>
      ATTORNEY/AGENT INFORMATION:
         NAME: POREMBSKI, PRISCILLA E.
         REGISTRATION NUMBER: 33,207
         REFERENCE/DOCKET NUMBER: 5527.PC.01
      TELECOMMUNICATION INFORMATION:
         TELEPHONE: 708-937-6365
         TELEFAX: 708-938-2623
   INFORMATION FOR SEQ ID NO: 11:
      SEQUENCE CHARACTERISTICS:
         LENGTH: 8912 base pairs
         TYPE: nucleic acid
         STRANDEDNESS: single
         TOPOLOGY: linear
      MOLECULE TYPE: DNA (genomic)
      SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-08-467-344A-11

Query Match   92.5%; Score 8692.6; DB 4; Length 8912;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 8781; Conservative 56; Mismatches 52; Indels 10; Gaps 8;
```

```
Qy     196 TGGGTTCGGTGGTGGTGGGACCGGGGACCGGGACCGGTCACTACCAAGGACGCAGACCTCTTTTGA
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11 TGGGTTCGGTGGTGGTGGGACCGGGGACCGGGACCGGTCACTACCAAGGACGCAGACCTCTTTTGA
```

```
Qy      71 GGGGGCACTGTAGGGAAGGACCGGGACCGGGACCGGTCACTACCAAGGACGCAGACCTCTTTTGA 130
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1

Qy     256 GGCGGCACTGTAGGGAAGGACCGGGGACCGGGACCGGTCACTACCAAGGACGCAGACCTCTTTTTGA 315

Db     316 GTATCACGGCCTCCGGAAGTAGTTGGGCAAGCCACCTATATGTTGGGATGGTTGGGT 375
           |:||||||||||||||||||||||||||||||||||:|:||:||||||||||||||
Qy     131 GTATCACGGCCTCCGGAAGTAGTTGGGCAAGCCACCCAYATGTGTTGGGATGGTTGGGT 190

Db     376 TAGCCATCCATACCGTACTGCCTGATAGGGTCCTGCGAGGGGATCTGGGAGTCTCGTAG 435
           |||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy     191 TAGCCATCCATACCGTACTGCCTGATAGGGTCCTGCGAGGGGATCTGGGAGTCTCGTAG 250

Db     436 ACCGTAGCAGACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACCTGGCGCCCAGAACGCG 495
           ||||||||||||||||||||||:|:|||||||||||||||||||||||:||:|||||||||||
Qy     251 ACCGTAGCAGACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACCTGCGRCCCAGAACGCG 310

Db     496 CAAGAACAAGCAGACGCAGCGTTCATATCCTGTGTCCATTAAAACATCTGTTGAAAGGGG 555
           |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Qy     311 CAAGAACAAGCAGACGCAGCGTTCATATCCTGTGTCCATTAAAACATCTGTTGAAAGGGG 370

Db     556 ACAACGAGCAAAGCGCAAAGTCCAGCGCGATGTCTGCGCCTCGTAATTACAAAATTGCTGG 615
           |||||||:|||||||||||||||||||||||||||||:||:|||||||||||||||||||
Qy     371 ACAACGAGCAARGCGCAAAGTCCAGCGCGATGTCTGCGCCTCGTAATTACAAAATTGCTGG 430

Db     616 TATTCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGCTCATGGTTGGGGACG 675
           |||||||||||||||||||||||||||||||||||||:|||||:||||||||||||||||
Qy     431 TATTCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTRCCAGCTCATGGTTGGGGACG 490

Db     676 CCAAGACCCTCGCCATAAGTCTCGCAATCTTCTGGAATCCTCTGGATTACCCTTTGGGGTG 735
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     491 CCAAGACCCTCGCCATAAGTCTCGCAATCTTCTGGAATCCTCTGGATTACCCTTTGGGGTG 550

Db     736 GATTGGTGATGTTACAACTCACACACCTTCAGTAGGCCCGTGGTGGCAGGAGCGGTCGT 795
           |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Qy     551 GATTGGTGATGTTACAACTCACACACCTTCAGTAGGCCCGTGGTGGCAGGAGCGGTCGT 610

Db     796 TCGACCAGTGCTGCCAGATAGTACGCTTGCTGGAGGATGGATGCAACTGGGCTACTGGTTG 855
           |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Qy     611 TCGACCAGTGCTGCCAGATAGTACGCTTGCTGGAGGATGGATGCAACTGGGCTACTGGTTG 670

Db     856 GTTCGGTGTCCACCTTTTTGTGGTATGTCGCTATCTTGGCCTGTCCCGTGTAGTGGGGC 915
           |||||||||||||||||||||||||||||:||:|||||||||||||||||||||||||||
Qy     671 GTTCGGTGTCGGTCCACCTTTTTGTGGTATGTCGCTGTCATYTTTGGCCTGTCCCGTGTAGTGGGGGC 730

Db     916 GCGGGTCACTGACCCAGACACAAATACCACACAATCCTGACCAATTGCTGCCAGCGTAATCA 975
           |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Qy     731 GCGGGTCACTGACCCAGACACAAATACCACAAATACCACCAATCCTGACCAATTGCTGCCAGCGTAATCA 790

Db     976 GGTTATCATTGTTCTCCTTCCACTTGCTACACGAGCCTGGTTGGTCGTGTGTCTGTCGGA 1035
           ||||||||:||||||:|:||||||||:|||||||:|||||||||||||::|:|||||||
Qy     791 GGTTATCAYTGTTCTCCTTCCACTTGCTACACGAGCCTGGTTGTGTGATCTGTGTGYGGA 850

Db    1036 CGAGTGCTGGGTTCCCGCCAATCCGTACATCTCACACCCTTCACACCCCTTCACACCTCCACGGA 1095
           ||||||||||||||||||||||||||||:||||:||||||||||||||||||||||||||
Qy     851 CGAGTGCTGGGTTCCCGCCAATCCRTACATCCRTACACTCTCACACCCTTCACACCTCCACGGA 910

Db    1096 CTCCTTCTTGGCTACCACATTGATTTTTGTTATGGGCGTCTTGTGACCTGTGACGCCCT 1155
           ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Qy     911 CTCCTTCTTGGCTACCACATTGATTTTTGTTATGGGCGTCTTGTGACCTGTGACGCCCT 970

Db    1156 TGACATTGGTGAGTTGTGGTGGTCGTGTATTAGTCGGTGTATTAGTCGGTGTTGTCAGGCACTG 1215
           ||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     971 TGACATTGGTGAGTTGTGGTGGTCGTGTATTAGTCGGTGTATTAGTCGGTGTTGTCAGGCACTG 1030

Db    1216 GCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGAAGTGCCCACTGGAAT 1275
           ||||||||||||||||||||||||:||||||:|||:|||||||||:|:|||||||||||
Qy    1031 GCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGAAKTGCCTACTGGAAT 1090

Db    1276 AGATCCTGGTTCCTAGGGTTTATCGGGTGGATGGCCGCAAGGTCGAGGCTGTCATCTT 1335
           ||||||||:|||||||||||||||||||||||:|||||||||||||||||||||||||
Qy    1091 AGATCCTGGTTCCTAGGGTTTATCGGGTGGATGGCCGCAAGGTCGAGGCTGTCATCTT 1150

Db    1336 CTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTATGTTTAGCAGTGTACA 1395
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1151 CTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTATGTTTAGCAGTGTACA 1215
```

```
Db  1151 CTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTATGTTTAGCAGTGTACA 1210

Qy  1396 CTACCTGGCGGTTGGGCGCTCTGATCTACTACTAC:||||||||||||:||||||||||| 1455
Db  1211 CTACCTGGCGGTTGGGCGCTCTGATCTACTACTAYGCCTCTCGGGGCAAGTGGTATCAGTTGCT 1270

Qy  1456 CCTAGCGCTTATGCTTACATAGAAGCGACCTCTGGAAACCCATCAGGGTGCCCACTGG 1515
    1456 ||||||||||:||||||||||||||||:||||||||||||||||||||
Db  1271 CCTAGCGCTTAYGCTTACATAGAAGCGACCTCYATCAGGGTGCCCACTGG 1330

Qy  1516 ATGCTTCAATAGCTGAGTTTGCTCGCTTTGATGATACCATGTCCTTGCCACTCTTATTT 1575
Db  1331 ATGCTTCAATAGCTGAGTTTGCTCGCCTTTGATGATACCATGTCCTTGCCACTCTTATTT 1390

Qy  1576 GAGTGAGAATGTGTCAGAAGTCATTTGTTACAGTCCAAAGTGACCAGGCCTATCACTCT 1635
Db  1391 GAGTGAGAATGTGTCAGAAGTCATTTGTTACAGTCGACAAAGTGGACCAGGCCTGTCACTCT 1450

Qy  1636 AGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGGTGCGAGGGGATGTAT 1695
Db  1451 AGAGTATAABAACTCCATATCTTGGTACCCCTATACAATCCCTGGTGCGAGGGGATGTAT 1510

Qy  1696 GGTTAAATTCAAAAATAACACAATGGGGTTGCTCGGTATTCGCAATGTGCCATCGTACTG 1755
Db  1511 GGTTAAATTCAAAAATAACACAATGGGGTTGCTCGCG-WWTCGCAATGTGCCATCGTACTG 1569

Qy  1756 CACTATGGGCACTGATGCAGTGGAACGACACTCGCAACACTTACGGACATGCGGTGT 1815
Db  1570 CACTATGGGCACTGATGCAGTGGAASSACAGTCGCAACACTTACGGACATGCGGTGT 1629

Qy  1816 AACACCATGCTAACAACCGATGCACAACCGATGCCTCAGCCCTGAAATTGGCTATATTACA 1875
Db  1630 AACACCATGCTAACAACCGATGCACAACCGATGCCTCAGCCCTGAAATTGGCTATATTACA 1689

Qy  1876 ATACCCTGGGTCTAAAAGAAATGTTTAAACCTTCATAATTTGGATGTCAGGCCATTTGTATTT 1935
Db  1690 ATACCCTGGGTCTAAAAGAAATGTTTAAACCTTCATAATTTGGATGTCAGGCCACTTGTATTT 1749

Qy  1936 TGAGGGATCAGATACCCCTATAGTTTACTTTATGACCCTGTGAATTCCACTCTCCTACC 1995
Db  1750 TGAGGGATCAGATAGTAGAATATATCAGGTTTACTTTATTGACCCTGTGAATTCCACTCTCCTACC 1809

Qy  1996 ACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCTGTTGGTACGGTGTTCTTGGTTACA 2055
Db  1810 ACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCTGTTGGTACGGTGTTCTTGGTTACA 1869

Qy  2056 GGTTCCGCAAGGGTTTTACAGTTTTTACAGCCGGTGATGTGAAAAGACCTAGCCACCAGGATTGATCACCAAAGA 2115
Db  1870 GGTTCCGCAA-GGTTTTACAGTTTTTACAGCCGGTGATGTGAAAAGACCTAGCCACCAGGATTGATCACCAAAGA 1928

Qy  2116 CAAAGGCCTGGAAAAATTATCAGGTCTTATATATTCCCGCACGGGTGCTTTGTCTCTTACGGG 2175
Db  1929 CAAAGCCTGGAAAAATTATCAGYTCTTATATATTCCCGCACGGGTGCTTTGTCTCTTACGGG 1988

Qy  2176 AGTTACCACCAAGGGGCGTGGTCTGATGATGTGTTGTGGCAGCAAGTATCTTAT 2235
Db  1989 GGTTCCGCAA-GGTTTTACAGTTTTTGTTTGTGGCAGCAAGTATCTTAT 2048

Qy  2236 TTTAGGCCTACCCTCTGTTACTTGTCCCCTTGTTTGGGCGCGGCTTCTGTTTGCG 2295
Db  2049 TTTAGGCCTACCCTCTGTTACTTGTCCCCTTGTTTGGGCGCGGCTTCTGTTTGCG 2108

Qy  2296 TCCTGTGCTCTCCATCCCAGTCGTATCTCCAAGCTGGCTGGATGTTTGTGCTAAAGCTCA 2355
Db  2109 TCCTGTGCTCTCCATCCCAGTCGTATCTCCAAGCTGGCTGGATGTTTGTGCTAAAGCTCA 2168

Qy  2356 AGTAGCTCCTTTTGCTTTTGATTTTCTTCATCTGTTGCTATCTCCGCTGCTATCTGTTTGCMCTTTACGTTA 2415
Db  2169 AGTAGCTCMTTTTGCTTTTGATTTTCTTCATCTGTTGCTATCTCCGCTGCTATCTGTTTGCMCTTTACGTTA 2228

Qy  2416 TGCTGCCCTTTTAGGGTTTGTGCCCATGGCTTGCCCTAACTTTCTTTTGTTGC 2475
Db  2229 TGCTGCCCTTTTAGGGTTTGTGCCCATGGCTTGCCCTAACTTTCTTTTGTTGC 2288
```

```
Qy  2476 AGCAGCTGCTCGCCCAACCAGATTATGACTGGTGGTGGCTGACTGTCACA 2535
Db  2289 AGCAGCTGCTCGCCCAACCAGATTATGACTGGTGGTGGCTGACTGTCACA 2348

Qy  2536 TTTGTGGGCCGGCCGTAACCGTGGTCACCGCATAGCTCGCTTGTGTAGTCCTTGGCCCTCT 2595
Db  2349 TTTGTGGGCCGGCCGTGACCGTGGTCA-CGCATAGCTCGCTTGTGTAGTCCTTGGCCCTCT 2407

Qy  2596 GGTAGGCGTTT-TAACCCTCTTGCATTTGGTTACGCCTGCTTCAGCTTTTGATACCGGAGA 2654
Db  2408 GGTAGGCGTTTYTAACCCTCTTGCATTTSSTKACGCCCTGCTT-AGCTTTTGACACCGGAGA 2466

Qy  2655 TAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCATGTCTCGTTTTGGCT 2714
Db  2467 TAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCATGTCTCGTTTTGGCT 2526

Qy  2715 TCTTTGCTCACTTGTTACCTCGGCTGCGTGTTTAGTTAACTCCTATCTTGGCAACGTTGGG 2774
Db  2527 TCTTTGCTCACTTGTTACCTCGCTGCGTGTTTAGTTAACTCCTATCTTTGGCAACGTTGGG 2586

Qy  2775 AGAATTGGTTTTGGAACGTTACACTAGACCGGAGAGGTTTTTCCTTGTGTCGTGTTTGTT 2834
Db  2587 AGAATTGGTTTTGGAACGTTACACTAAGACCGGAGAGGTTTTCCTTGYGCTGTGTTTGVT 2646

Qy  2835 TCCCCGGTGCGACATATGACGGCGCTGGTGTGACTTTCTGTGTGTGTCACGTAGCTCTTCTAT 2894
Db  2647 TCCCCGGTGCGACATATGACGTGCTGGTGTGACWTTCTGTGTGTGTCACGTAGCTCTTCTAT 2706

Qy  2895 GTTTAACATCCAGTGCGAGCATCGTTCTTGGGACTCTAGGGTTAGGGCCCATAGAA 2954
Db  2707 GTTTAACATCCAGTGCGAGCATCGTTCTTGGGACTCTAGGGTTAGGGCCCATAGAA 2766

Qy  2955 TGTTGGTGCGTCTCGGAAAAGTGTCATGCTTGCAGGCCTATTCTCATTATGTTCTTAAGTTTTTCC 3014
Db  2767 TGTTGGTGCGTCTCGGAAAAGTGTCATGCTTGCAGGCCTATTCTCATTATGTTCTTAAGTTTTTCC 2826

Qy  3015 TCTTAGTGTTGGTGGAGAATGGTGTGTGTTTTTCTATAAGCACTTGCATGGTGATGTCTTGC 3074
Db  2827 TCTTAGTGTTGGTGGAGAATGGTGTGTGTTTTTCTAKAAGCACTTGCATGGTGATGTCTTGC 2886

Qy  3075 CTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCATTTTTCCCTTTGAAGGCAAGG 3134
Db  2887 CTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCATTTTTCCCTTTGAAGGCAAGG 2946

Qy  3135 CAAGGGTCTATAGGAATGAAGGAAGACGCTTCGGCGTGTGGGGCACCGGTTGATGGTTTGC 3194
Db  2947 CAAGGGTCTATAGGAATGAAGGAAGACGCTTGGSKKGTGGGGCACCGGTTGATGGTTTGS 3006

Qy  3195 CCGTTGTTGCGCGTCTCGGCGACCTTGTTTTCGCAGGGTTGGCTGTATGCCGCACGCGCATGGGT 3254
Db  3007 SCGTTGTTBGCGCGTCTCGGCGACCTTGTTTTCGCAGGGTTAGCTGTATGCCGCACGCGCATGGGT 3066

Qy  3255 GGGCCATTACCGGACCACCTTTTACGCTGGATGGGTTGTGTGGCAAGTATCGTCAGGCGA 3314
Db  3067 GGGCCATTACCGGACCACCTTTTACGCTGGATGGGTTGTGTGGCAAGTATCGTCAGGCA 3126

Qy  3315 TGGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAACTATCTTCAGATTAG 3374
Db  3127 TGGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAACTATCTTCAGATTAG 3186

Qy  3375 GATCTCTGGCCACTAGCTACATGGGGATTGTTTGTGACAACGTGTTGTACTGCTCACC 3434
Db  3187 GATCTCTGGCCACTAGCTACATGGGGATTGTTTGTGACAACGTGTTGAATTACTGCTCACC 3246

Qy  3435 ATGGCAGCAAGGGCGGCGGTTGGCTCTATACCCCAATAACCGTTG 3494
Db  3247 ATGGCAGCAACGGGGGCCGGGTTGGCTCCATCACCACACCCAATAACCGTTG 3306

Qy  3495 ACGCGGCTAATGACCAGGACATCTCATCAACCACACCATGGGAGCTGGGAGCTGGGTTCCCTTACTCCGT 3554
Db  3307 ACGCGGCTAATGACCAGGACATCTATCAACCACCCATGTGGAGCTGGGTTCCCTTACTCCGT 3366
```

Mon Oct 31 10:27:43, 2005   us-10-009-002-1.rni   Page 55

```
Db  5527  TGACATTTAAATGCTTGACTGGTGAGTGGCYCACTATGGATCAGCTTGCTGGTTAGTCT  5586

Qy  5775  ACTCCGCGTTCAATCCGGCCACCGGGCCCAGGAGTTGTGGGCGTCTTGTTGTCAATGTTTG  5834
          ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db  5587  ACTCCGCGTTCAATCCGGCCACCGGGCCCAGGAGTTGTGGGCGTCTTGTTGTCAATGTTTG  5646

Qy  5835  CTTTGACAACAGCAGGGCCAGGGCCAGATCACTCGGCCCAACAGACTTCTTACTATGCTTACTAGGA  5894
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5647  CTTTGACAACAGCAGGGCCAGGGCCAGATCACTCGGCCCAACAGACTTCTTACTATGCTTGCTAGGA  5706

Qy  5895  GCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCCGCAGGAAGATACTGG  5954
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  5707  GCAACACTGTATGTAARTGAGTACTTTATTGCCACTCGTGACATCCGCAGGAAGATACTGG  5766

Qy  5955  GCATTCTGGAGGCATCTACCCCCTGAGTGTCATATCAGCTGCATCCGTTGGCTCCACA  6014
          ||||||||||||||||||||||||||||| |||| ||| ||||||||||
Db  5767  GCATTCTGGAGGCATCTACCCCCTGAGTGTRTCATATCAGCTGCATCCGTTGGCTYCACA  5826

Qy  6015  CCCCGACGGAGGATGATTGCGGCGCCTCATTGGGGTCTAGAGATTTGGCAGTAGTGTGT  6074
          |||||||||||||||||||||||||||| ||||||||||||||||||
Db  5827  CCCCGACGGAGGATGATTGCGGCGCCTCATTGGGGTCTARAGATTTGGCAGTATGTGT  5886

Qy  6075  GCAATTTCTTTGTGATTTGCTTTAAATGTCCTTAAAGCTGGAGTTCAGAGCATGTTAACA  6134
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5887  GCAATTTCTTTGTGATTTGCTTTAAATGTCCTTAAAGCTGGAGTTCAGAGCATGTGTTAACA  5946

Qy  6135  TTCCTGGTTGTCCTTTTCTACAGCTGCCAGAGAGGGTACAAGGGCCCCTGGATTGGDATCAG  6194
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  5947  TTCCTGGTTGTCCTTTCTACAGCTGCCAGAGAGGGTACAAGGGCCCCTGGATTGGDATCAG  6006

Qy  6195  GTATGCTACACCGTGTCCATGCGGTGCTGAACTCATCTTTTCTGTTGAGAATGGTT  6254
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6007  GTATGCTACACCGTGTCCATGCGGTGCTGAACTCATCTTTTCTGTTGAGAATGGTT  6066

Qy  6255  TTGCAAAACTTTACAAAGGAGACACCCAGACTTGTCAAATTACTGGAGGGCTGTTCCAG  6314
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6067  TTGCAAAACTTTACAAAGGAGACACCCAGACTTGTCAAATTACTGGAGGGCTGTTCCAG  6126

Qy  6315  TCAACGCTAGGCGTGTGGGTCGGCTAGACCGGACCCAACTGATTGGACTAGTCTTGTCTG  6374
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6127  TCAACGCTAGGCGTGTGGGTCGGCTAGACCGGACCCAACTGATTGGACTAGTCTTGTCG  6186

Qy  6375  TCAATTATGGCGTTAGGGACTAGTGTAAATGGAGAAAATGAGGAGATCACATTTTTGTTA  6434
          |||||||||||||||||||||||||||||||||||||||||||||
Db  6187  TCAATTATGGCGTTAGGGACTAGTGTAAATGAGGAGAAAATGGGGAGATCACATTTTTGTTA  6246

Qy  6435  CAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGTGCCCCCAACCTTGAGGAGTTGCAG  6494
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  6247  CAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGTGCCCCCAACCTTGAGGAGTTGCAG  6306

Qy  6495  TGGCCGGTGGACGGGCTACAGGTTCAGGTTTATCTAGGTGAGCCCAAAACTCCTTGGACGA  6554
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6307  TGGCCGGTGGACGGGCTACAGGTTCAGGTTTATCTAGGTGAGCCCAAAACTCCTTGGACGA  6366

Qy  6555  CATCTGCTTGCTGTTACGGTTCGCCGATGCGCATGCAACTTAATTTGCGTGATGCACTTGA  6614
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  6367  CATCTGCTTGCTGTTACGGTTCGCCGATGCGCATGCAACTTAATTTGCGTGATGCACTTGA  6426

Qy  6615  TTGACGGTCACACACCTGGTCGCCATGGTGCCACAAACACACTCCTAGTGATTGCTTGACACT  6674
          ||||||||||||||||||||||||||||||||||||||||
Db  6427  TTGACGGACACACACCTGGTCGCCATGGTGCCACAAACACACTCCTAGTGATTGCTTGAGCAA  6486

Qy  6675  ATGACAGGAGTTGCGCGGTACAAACACTCCTAGTTGCTTGATGAAGCCGCAGTGTCCGCTCTTGTTT  6734
          ||||||||||||||||||||||||||||||||||||||||
Db  6487  ATGACAGGAGTTGCGCGGTACAAACACTCCTAGTTGCTTGATGAAGCCGCAGTGTCCGCTCTTGTTT  6546

Qy  6735  TCAAACAGGAGTTGCGCGGTACAAATTGCTTGAGGCAATTTCAGCTTGCGGTTGACA  6794
          ||||||||||||||||||||||||||||||||||||||||
Db  6547  TCAAACAGGAGTTGCGCGGTACAAATTGCTTGAGGCAATTTCAGCTTGCGGTTGACA  6606

Qy  6795  CCACCAAACTGCCAGCCCCCTCC--ATCGAAGAGGTAGTGTAAGAAAGGCGCCAGTTCC  6851
          ||||||||||||||||||||||||||||||||||||||||
Db  6607  CCACCAAACTGCCAGCCCCCTCCCCCAGATCGAAGAGGTAGTGTAAGAAAGGCGCCAGTTCC  6666
```

```
Qy  6852  GGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCCTCCGAGATCCGTCCCCAGGAGTGTCAT  6911
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  6667  GGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCCTCCGAGATCCGTCCCCAGGAGTGTCAT  6726

Qy  6912  GTCCTGAAAGCTGCAACGAACGAAGTGACCGTTAGAAGGTCCTTCAAACCTCCCTTCAC  6971
          |||||||||||||||||||||||||||||  :  | : | |:|||||||||||||
Db  6727  GTCCTGAAAGCTGCAACGAACGAAGTGACCGTTAGAAGGTCCTTCAAMCCTCCCTTCTTCAC  6786

Qy  6972  CACCTGTTCTACAGTTGGCCACTGCCAGGAAGATCCCGCAGGGTGAGTGTAACCCTT  7031
          ||||||||: ||| | |||||||||||||||||||||||||||||
Db  6787  CACCTGTTCTRCAGTTGGCCACTGCCAGGGTGGGAGGCAGGTGAGTGTAACCCTT  6846

Qy  7032  TCACTGCAATTGGATGTGCAATGACCGAAACAGGCGGAGGCCCTGATGATTTACCCAGTT  7091
          |||||||||||||||||||||||||||||||||||   | |||||||||||
Db  6847  TCACTGCAATTGGATGTGCAATGACCGAAACAARGYGGAGCCCCWMAKRATTTACCCAGTT  6906

Qy  7092  ACCCTCCCAAAAAGGAGGTCTCTGAAATGGTCAGACGAAGATTGGTGCGACGGCTACAACCG  7151
          ||||||||||||||||||||||||||||||||||||||||
Db  6907  ACCCTCCCAAAAAGGAGGTCTCTGAAATGGTCAGACGAAGATTGGTCAACGACTACAACCG  6966

Qy  7152  CTTCCAGCTACGTTACTGGCCCCCGTACCCTAAGATACGGGAAAGGATTCCACTCAGT  7211
          ||||||||||||||||||||||||||||||||||||||||
Db  6967  CTTCCAGCTACGTTACTGGCCCCCGTACCCTAAGATACGGGCAAGGATTCCACTCAAT  7026

Qy  7212  CAGCCCCGCCAAAACGGCTACACAAAAAGAGTGGAAAGAGTGAGTTTCGTGCAGCA  7271
          ||||||||||||||||||||||||||||||||||||||||
Db  7027  CAGCCCACCGCCAAAACGGCTACACAAAAAGAGTGGAAAGAGTGAGTTTCGTGCAGCA  7086

Qy  7272  TGAGCTACACCGTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAGTTCTGTCTGCAA  7331
          |||||||||||||||||||||||||||||||||||||||||
Db  7087  TGAGCTACACCGTGGACCGACGTGATTAGCTTCTAAAACTGCTTCTAAGTTCTGTCTGCAA  7146

Qy  7332  CTCGGGCCATCACTAGTGGTTCCTCAAACAAGATCATTGGTGTATGTGACTGAGCCGC  7391
          ||||||||||||||||||||||||||||||||||||||||
Db  7147  CTCGGGCCATCACTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGTGACTGAGCCGC  7206

Qy  7392  GGGATGCGGAGTTAGAAGAACAAAAACGTCACTATTAATAGACAACCTCTGTTCCCCCAT  7451
          ||||||||||||||||||||||||||||||||||||||||
Db  7207  GGGATGCGGAGTTAGAAGAACAAAAACGTCACTATTAATAGACAACCTCTGTTCCCCCAT  7266

Qy  7452  CATACCACAAGCAAGTGGCTAAGGAAAAGCTTCAAAAGTTGTCGGTGTCATGT  7511
          ||||||||||||||||||||||||||||||||||||||||
Db  7267  CATACCACAAGCAAGTGGCTAAGGAAAAGCTTCAAAAGTTGTCGGTGTCATGT  7326

Qy  7512  GGGACTATGATGAAGTAGCAGCTCACACGCCCTCTTAGTCTGCTCAAGTCCCACATCACTG  7571
          ||||||||||||||||||||||||||||||||||||||||
Db  7327  GGGACTATGATGAAGTAGCAGCTCACACGCCCTCTTAGTCTGCTCAAGTCCCACATCACTG  7386

Qy  7572  GCCTTCGGGCACTGATGTTCGTTCTGGACTTGCAGCAGCCCCAAGGCGTGTTCTGGACTTGCAGA  7631
          ||||||||||||||||||||||||||||||||||||||||
Db  7387  GCCTTCGGGCACTGATGTTCGTTCTGGACTTGCAGCAGCCCCAAGGCGTGTTCTGGACTTGCAGA  7446

Qy  7632  AGTGTGTCGAGGCAGGTGAGATCATTATCGGCAACTGTGATAGTTCCAAAGG  7691
          ||||||||||||||||||||||||||||||||||||||||
Db  7447  AGTGTGTCGAGGCAGGTGAGATCATTATCGGCAACTGTGATAGTTCCAAAGG  7506

Qy  7692  AGGAGGTCTTCGTGAGACAGCCCCCAGAGAAACAACAAAGAAACCCCAAGGCTTATCTCGT  7751
          ||||||||||||||||||||||||||||||||||||||||
Db  7507  AGGAGGTCTTCGTGAGACAGCCCCCAGAGAAACAACAAAGAAACCCCAAGGCTTATCTCGT  7566

Qy  7752  ACCCCACCTTGAAATGAGAGATGTGTTGAGAAGATGTTGAGAGATGAGAGTACTACGGTCCTGACG  7811
          ||||||||||||||||||||||||||||||||||||||||
Db  7567  ACCCCACCTTGAAATGAGAGATGTGTTGAGAGATGTTGAGAGATACTACGGTCCTGACG  7626

Qy  7812  TAGTTAAAGCTGTCATGGGAGATCGCGGTTTGTAGATCCACGTTGTAGATCCGTGTCAAGC  7871
          ||||||||||||||||||||||||||||||||||||||||
Db  7627  TAGTTAAAGCTGTCATGGGAGATCGCGGTTTGTAGATCCACGTTGTAGATCCGTGTCAAGC  7686

Qy  7872  GTCTGTTTGTCGATGTGGTCACCCGAGCCACATGCGATCGCGAGCCACATGCGATACAGTGTGTTTTG  7931
          ||||||||||||||||||||||||||||||||||||||||
Db  7687  GTCTGTTTGTCGATGTGGTCACCCGAGCCAGCAGTCGGATCGCGAGCCACATGCGATACAGTGTGTTTTG  7746
```

Db   8825 TCATTGCTGTTGGACTAGCCATCAGCTGAACCCCAAATTCAAAATTAATTAACAGTTTT 8884
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   9072 TTTTTTTTTTTTTTTTT 9090
            |||||||||||||||||
Db   8885 TTTTTTTTTTTTTTTTT 8903

RESULT 13
US-08-424-550B-11

; Sequence 11, Application US/08424550B
; Patent No. 672016б
; GENERAL INFORMATION:
;   APPLICANT: JOHN N. SIMONS
;   APPLICANT: TAMI J. PILOT-MATIAS
;   APPLICANT: GEORGE J. DAWSON
;   APPLICANT: GEORGE G. SCHLAUDER
;   APPLICANT: SURESH M. DESAI
;   APPLICANT: THOMAS P. LEARY
;   APPLICANT: ANTHONY SCOTT MUERHOFF
;   APPLICANT: JAMES C. ERKER
;   APPLICANT: SHERI L. BUIJK
;   APPLICANT: ISA K. MUSHAHWAR
;   TITLE OF INVENTION: NON-A, NON-B. NON-C, NON-D, NON-E HEPATITIS
;   TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
;   NUMBER OF SEQUENCES: 716
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: ABBOTT LABORATORIES   D377/AP6D
;   STREET: 100 ABBOTT PARK ROAD
;   CITY: ABBOTT PARK
;   STATE: IL
;   COUNTRY: USA
;   ZIP: 60064-3500
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/424,550B
;   FILING DATE:
;   CLASSIFICATION: 435435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: POREMBSKI, PRISCILLA E.
;   REGISTRATION NUMBER: 33,207
;   REFERENCE/DOCKET NUMBER: 5527.PC.01
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 708-937-6365
;   TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 11:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 8912 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-08-424-550B-11

Query Match   92.5%; Score 8692.6; DB 4; Length 8912;
Best Local Similarity   98.7%; Pred. No. 0;
Matches 8781; Conservative 56; Mismatches 52; Indels 10; Gaps 8;

Qy   196 TGGGTTCGGTGGTGGTGGTGCGGCTTTAGGCACCGGCCTCCACGCCGCCACCACCTCCAGATAGAGC 255
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11 TGGGTTCGGTGGTGGTGGTGCGGCTTTAGGCACCGGCCTCCACGCCGCCACCACCTCCAGATAGAGC   70

Qy   256 GGGCGGCACTGTAGGGAAGACCGGGGACCGGTCTACCAAGGACACCGCAGACCTCTTTTTGA 315
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   71 GGGCGGCACTGTAGGGAAGACCGGGGATCCGGTCTACCAAGGACACCGCAGACCTCTTTTTGA 130

Qy   316 GTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTTGGGATGGTTGGGGT 375
            |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   131 GTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACCTAYATGTTGGGATGGTTGGGGT 190

Qy   7932 ACAGTACCATCACACCCGAGATCATGGTGGAGACAGACATCTACTCAGCAGCTAAAC 7991
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7747 ACAGTACCATCACACCCGAGATCATGGTGGAGACAGACATCTACTCAGCAGCTAAAC 7806

Qy   7992 TCAGTGACCAACACCGAGCTGGCATTCACACACCATTGCGAGGCAGTTATACGCTGGAGGAC 8051
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7807 TCAGTGACCAACACCGAGCTGGCATTCACACACCATTGCGAGGCAGTTATACGCTGGAGGAC 7866

Qy   8052 CGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTGAGGTCTTCCGGCGTCT 8111
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7867 CGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTGAGGTCTTCCGGCGTCT 7926

Qy   8112 ATACTACCTCAAGTTCAACAGTTGACCTGCTGGCTCAAGAGTAAATGCTGCAGCCGAAC 8171
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7927 ATACTACCTCAAGTTCAACAGTTGACCTGCTGGCTCAAGAGTAAATGCTGCAGCCGAAC 7986

Qy   8172 AGGCTGGCATGAAGAACCCTCGCTTCCTTATTTGCGGCGATGATTGCACCGTAATTTGGA 8231
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7987 AGGCTGGCATGAAGAACCCTCGCTTCCTTATTTGCGGCGATGATTGCACCGTAATTTGA 8046

Qy   8232 AGAGCGCCGGAGCGAGATGCAGACAAACAAGCAATGCGTGTCTTGCTAGTCGATGAAGG 8291
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8047 AGAGCGCCGGAGCGAGATGCAGACAAACAAGCAATGCGTGTCTTGCTAGTCGATGAGAAGG 8106

Qy   8292 TGATGGGTGCACCACAGAGTTGTGCGCTCAAGACCAAATACAGTTTGGAAGAATTAACAT 8351
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8107 TGATGGGTGCACCACAGAGTTGTGCGCTCAAGACCAAATACAGTTTGGAAGAATTAACAT 8166

Qy   8352 CATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTACTACTTTCTA 8411
            |||:|||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db   8167 CATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTACTACTTTCTTA 8226

Qy   8412 CAAGAGATCCTCGTATCCCCCCCTTGGCAGGTGCTCTGCGAGGGTCTGGGATACAACCCCA 8471
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8227 CAAGAGATCCTCGTATCCCCCCCTTGGCGAGGTGCTCTGCGAGGGTCTGGGATACAACCCCA 8286

Qy   8472 GTGCTGCCGTGGATTGGGTATCTAATAACATCACTACCCCATGTTTGTGGGTTAGCCGTGTGT 8531
            |:::|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8287 GKGCKGCCGTGGATTGGGTATCTAATAACATCACTACCCCATGTTTGTGGGTTAGCCGTGTGT 8346

Qy   8532 TGGCTGTGTCCATTTCATGGAGCAGATGCGTCTCTTTGAGGACAAACTTCCCGAGACTGTGACCT 8591
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8347 TGGCTGTGTCCATTTCATGGAGCAGATGCGTCTCTTTGAGGACAAACTTCCCGAGACTGTGACCT 8406

Qy   8592 TTGACTGGTATGGGAAAAATTATACGGTCGTAGAGAGATCTGCCCGCCAGCATCATTGCTG 8651
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8407 TTGACTGGTATGGGAAAAATTATACGGTCGTAGAGAGATCTGCCCGCCAGCATCATTGCTG 8466

Qy   8652 GTGTGCACGGTATTGAGGCTTTCTCGGTGGTGCGCTACACCAACGCTGAGATCCTCAGAG 8711
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8467 GTGTGCACGGTATTGAGGCTTTCTCGGTGGTGCGCTACACCAACGCTGAGATCCTCAGAG 8526

Qy   8712 TTTCCCAATCACTAACAGACACGACCATGACCATGCGCCCCCCCTGCGACCTGCGACGAAAGAAAGCCA 8771
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8527 TTTCCCAATCACTAACAGACACGACCATGACCATGCGCCCCCCCTGCGACCTGCGACGAAAGAAAGCCA 8586

Qy   8772 GGGCCGGTCCTGCCGCCCTGGCCCAGCCCCAGAGGCGTGGGCCGAGCACACGCAAAATTGGCTCGCTTCC 8831
            ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   8587 GGGCCGGTCCTGCCGCCCTGGCCCAGCCCCAGAGGCGTGGGCCGAGCACACG-AAAATTGGCTCGCTTCC 8645

Qy   8832 TTCTCTGCATGCTACATCTAGACCTCTACAGATTTGATAAGACGAGCGTTGGCGTCGCTCGT 8891
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8646 TTCTCTGCATGCTACATCTAGACCTCTACAGATTTGATAAGACGAGCGTTGGCGTTGCTCGT 8705

Qy   8892 ACACCACTTTCAATTATTGTGATGTGTTATTCACCAC 8951
            |||||||||||||||||||||||||||||||||||::
Db   8706 ACACCACTTTCAATTATTGTGATGTTTACT-CCCSGAGRGGGGATGTGTTTATTCACCAC 8764

Qy   8952 AGAGAAGATTGCAGAAGTTCCTTGTGAAGTATTGGCTGTCATTGTTTTTGCCCTAGGGC 9011
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8765 AGAGAAGATTGCAGAAGTTTCTTGTGAAGTATTTGGCTGTCATTGTTTTTGCCCTAGGGC 8824

Qy   9012 TCATTGCTGTTGGATTAGCCATCAGCTGAACCCCAAATTCAAAATTAACTAACAGTTTT 9071
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8825 TCATTGCTGTTGGACTAGCCATCAGCTGAACCCCAAATTCAAAATTAATTAACAGTTTT 8884

```
Qy  1456  CCTAGCGCTTATGCTTTACATAGAAGCGACCTCTGGAAACCCATCAGGGTGCCCACTGG  1515
          |||||||||||:|||||||||||||||| |||||||||||||||||:|||||||||||||
Db  1271  CCTAGCGCTTAYGCTTTACATAGAAGGGACCTCTGGAAACCCYATCAGGGTGCCCACTGG  1330

Qy  1516  ATGCTCAATAGCTGAGTTTGCTCGCCTTGATGATGACCATGTCCTTGCCACTCTTATTT  1575
          |||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db  1331  ATGCTCAATAGCTGAGTTTGCTCGCCTTGATGATGACCATGTCCTTGCCATCTCTTATTT  1390

Qy  1576  GAGTGAGAATGTGTCAGAGTCATTGTTACAGTCAAAGTGGACCAGGCCTGTCACTCT  1635
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1391  GAGTGAGAATGTGTCAGAGTCATTGTTACAGTCCAAAGTGGACCAGGCCTGTCACTCT  1450

Qy  1636  AGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCTGGTGCGGAGGGGATGTAT  1695
          |||||||||:|| :|||||||||||||||||||  ||||||||||||||||||||||||||
Db  1451  AGAGTATAABAACTCCATATCTTGGTACCCCTATACAATCCCTGGTGCGGAGGGGATGTAT  1510

Qy  1696  GGTTAAATTCAAAAATAACACACATGGGGTGCTCGCCGTATTCGGCAATGTGCCATCGTACTG  1755
          |||||||||||||||||||||||||||||||||||||||||| ::|||||||||||||||||
Db  1511  GGTTAAATTCAAAAATAACACACATGGGGTGCTCGCCGTCGG-WWTCGCAATGTGCCATCGTACTG  1569

Qy  1756  CACTATGGCACTGATGCAGTGTGGAACGACACTCGCAACACTTACGAAGCATGCGGTGT  1815
          ||||||||||||||||||||||||||||||||||| ||::|||||||||||||||||||
Db  1570  CACTATGGCACTGATGCAGTGTGGAASSACAGTCGCAACACTTACGAAGCATGCGGTGT  1629

Qy  1816  AACACCATGGCTAACAACCGATGGCACAACGGCTCAGCCCTGAAATTGGCTATATTACA  1875
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1630  AACACCATGGCTAACAACCGATGGCACAACGGCTCAGCCCTGAAATTGGCTATATTTACA  1689

Qy  1876  ATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTCAGGCCATTTGTATTT  1935
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1690  ATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTCAGGCCACTTGTATTT  1749

Qy  1936  TGAGGGATCAGATACCCCTATAGTTTACTTTTATGACCCGTGAATTCCTACCTCCTACC  1995
          |||||||||||||||||||||||||||||||||::|||||||||||||||||||||||||
Db  1750  TGAGGGATCAGATACCCCTATAGTTTACTTTACTTTATGACCCGTGAATTCCTACCTCCTACC  1809

Qy  1996  ACCGGAGAGTGGGCTAGGTTACAGTGTTGCCCGGTACCCCACCTGTGGTACGTGGTTACA  2055
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1810  ACCGGAGAGTGGGCTAGGTTGCCCGGTACCCCACCTGTGGTACGTGGTTCTTGGTTACA  1869

Qy  2056  GGTTCCGCAAGGGTTTACAGTGATGATGTGAAAGACCTAGCACAGGATTGATGACCAAAGA  2115
          |||||||||||  |||||||||||||||||||||||||||||||||||||||||||||||
Db  1870  GGTTCCGCAA-GTTTTACAGTGATGTGAAAGACCTAGCACAGGATTGATGACCAAAGA  1928

Qy  2116  CAAAGCCTGGAAAAATTATCAGTCTTATATTCCGACGGGTGCTTTGTCTCTTACGGG  2175
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1929  CAAAGCCTGGAAAAATTATCAGYTCTTATATTCCGACGGGTGCTTTGTGTCTTACGGG  1988

Qy  2176  AGTTACCACCAAGGCCGTGGTGCTAATTCTGTTGGGCAGCAAGTATCTTAT  2235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1989  AGTTACCACCAAGGCCGTGGTGCTAATTCTGTTGGGCAGCAAGTATCTTAT  2048

Qy  2236  TTTAGCCTACCTCTGTACTTGTCCCTTGTTTGGGCGCGCTTCGGTTACCCTTTGCG  2295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2049  TTTAGCCTACCTCTGTACTTGTCCCTTGTTTGTTTGGGCGCGCTTCGTTACMCTTTGCG  2108

Qy  2296  TCCTGTGTCCCATCCCAGTCGTATCTCCCAGTCGTATCTCGGGGATGTTTTGTCTAAAGCTCA  2355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2109  TCCTGTGTCCCATCCCAGTCGTATCTCCCAGTCGTATCTTTTGTCTAAAGCTCA  2168

Qy  2356  AGTAGCTCCTTTTGCTTGATTTTCTTCATCTGTTGCTATCTCCGCTGCAGGCTACGTTA  2415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2169  AGTAGCTCMTTTTGCTTTGATTTTCTTCATCTCCGCTGCAGGCTACGTTA  2228

Qy  2416  TGCTGCCCTTTAGGGTTGTGCCCATGGCTTGCCCCTAACTTTCTTTGTTGC  2475
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2229  TGCTGCCCTTTAGGGTTTGTGCCCATGGCTTGCCCCTAACTTTCTTTGTTGC  2288

Qy  2476  AGCAGCTGCTGCCCAACCAGATTATGACTGGTGGTCAGTGTAGT  2535
          ||||||||||||||||||||||||||||||||||||||||||||
Db  2289  AGCAGCTGCTGCCCAACCAGATTATGACTGGTGGTCAGTGTAGT  2348

Qy  2536  TTTGTGGGCCGGCCCGTAACCGTGGTCACCCGTCGCTTCGCTTGCCTTCT  2595
```

```
Qy  376   TAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCGGAGGGATCTGGGGAGTCTCGTAG  435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  191   TAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCGGAGGGATCTGGGGAGTCTCGTAG  250

Qy  436   ACCGTAGCACATGCCTGTTATTCTACTCAAACAAGTCCTGTACCTGCGCCCAGAACGCG  495
          |||||||||||||||||||||||||||||||||||||:||||||| ::||||||||||||
Db  251   ACCGTAGCACATGCCTGTTATTCTACTCAAACAAGTCCTGTACCTGCRCCCAGAACGCG  310

Qy  496   CAAGAACAAGCAGCAGCGCAGCGGCCTCATATCGTGTCCATTAAAACATCTGTTGAAAGGGG  555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  311   CAAGAACAAGCAGCAGCGCAGCGGCCTCATATCGTGTCCATTAAAACATCTGTTGAAAGGGG  370

Qy  556   ACAACGAGCAAAGCGCAAAGTCCAGCCGATGCTCGCCTCGGTACCCTGTAATTACAAAATTGCTGG  615
          ||||||||||||| ||||||||||||||||||||||||||||| |:|||||||||||||||
Db  371   ACAACGAGCAARGCGCAAAGTCCAGCCGATGCTCGCCTCGGTAATTACAAAATTGCTGG  430

Qy  616   TATCCATGATGGCTTGCAGACATTGGCTCAGGTGCTTGCCAGCTCATGGTTGGGGACG  675
          ||||||||||||||||||||||||||||||||||||||||||||: :|||||||||||
Db  431   TATCCATGATGGCTTGCAGACATTGGCTCAGGTGCTTGCCAGCTCATGGTTGGGGACG  490

Qy  676   CCAAGAACCCTGCGCCATAAGTCTCGCAATCTTCTGGAATCCTTCTGGGGTG  735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  491   CCAAGAACCCTGCGCCATAAGTCTCGCAATCTTCTGGAATCCTTCTGGGGGTG  550

Qy  736   GATTGGTGATGTTACAACTCACACACCTCCAGTAGGCCCGGTGGTGGCAGGAGCGGTGTCGT  795
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  551   GATTGGTGATGTTACAACTCACACACCTCCAGTAGGCCCGGTGGTGGCAGGAGCGGTGTCGT  610

Qy  796   TCGACCAGTCTGCCAGATAGTACGCTTGCTCGCTGGAGGATGGGAGTCAACTGGGCTACTCGTTG  855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  611   TCGACCAGTCTGCCAGATAGTACGCTTGCTCGCTGGAGGATGGGAGTCAACTGGGCTACTCGTTG  670

Qy  856   GTTATCTATTGTCTCCGTTCCGCGACCAATTGCTGCCAGCGAACTGCCAGCGTAATCA  915
          |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db  671   GTTATCTAYTGTCTCCGTTCCGCGACCAATTGCTGCCAGCGTAATCA  730

Qy  916   GCGGGTCACTGACCCGAGACACAAATACCACAATCCGTACCTTGCCTCACGAGCCTGGCACGGA  975
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  731   GCGGGTCACTGACCCGAGACACAAATACCACAATCCRTACCTTGCCTCACGAGCCTGGCACGGA  790

Qy  976   GGTTATCTATTGTCTGCGTGTGTGATCGTGCGGA  1035
          |||||||:|||||||||||||||||:||||||
Db  791   GGTTATCTAYTGTCTGCGTGTGTGATCGTGTGYGGA  850

Qy  1036  CGAGTGCTGGGTTCCGCGTACATCTCACACACCTTCCAAATTGGACTGGCTGCTGGCACGGA  1095
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  851   CGAGTGCTGGGTTCCGCGTACATCCCRTACACCTTCCAAATTGGACTRGGCTGGCACGGA  910

Qy  1096  CTCCTTCTTGGCTGACCATTGATTTTGTTATTGTTATGGGGTTACTTGTTACCTGCTGGAAT  1155
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  911   CTCCTTCTTGGCTGACCATTGATTTTGTTATGGGGTGGTATGGCTGGAAKTGGAAT  970

Qy  1156  TGACATTGGTGAGTTGTGTGGTGTGGTGTATTAGTTAGTCGGTGACTGGCTTGTCAGGCACTG  1215
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  971   TGACATTGGTGAGTTGTGTGGTGTGGTGTATTAGTTAGTCGGTGACTGGCTTGTCAGGCACTG  1030

Qy  1216  GCTTATTCACATAGAACTCAATGAAACTGTACTTGTTACCTGGAAGTGCCCACTGGAAT  1275
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1031  GCTTATTCACATAGAAACTCAATGAAAKTGTACTTGTTACCTGGAAKTGCCTACTGGAAT  1090

Qy  1276  AGATCCTGGGTTCCTAGGGTTTATCGGGTGGATAGGCTGCAAGGTCGAGGCTGTCATCTT  1335
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1091  AGATCCTGGGTTCCTAGGGTTTATCGGGTGGATAGGCCGCAAGGTCGAGGCTGTCATCTT  1150

Qy  1336  CTTGACCAAACTGGCTTCACAGAACCATTGCGACTATGTTTAGCAGTGTACA  1395
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1151  CTTGACCAAACTGGCTTCACAGAACCATTGCGACTATGTTTAGCAGTGTACA  1210

Qy  1396  CTACCTTGGCGGTTGGCCCGTTTCCTGATCGAATCATGTCCCTTGGTATCGATCAGTTGCT  1455
          ||||||||||||||||||||::|||||||||||||||||||||||||||||||||||||||
Db  1211  CTACCTTGGCGGTTGGCCCGTTCCTGATCGAATCACTAYGCCTCTGGTATCGATCAGTTGCT  1270
```

Left column:

```
           |||||||||||||:| |||||||| |||||||||||  :  -||||||||||||||
Qy  2349   TTTGTGGGGCCGGCCGTGACCGTTGGTCA-CGCATAGCTTGCTTGCTTGGTCCTTGGCCTCT  2407

Db  2596   GGTAGCGCTTT-TAACCCTCTTGCATTGGTTACGCCTGCTTCAGCTTTGATACCGAGA  2654
           |||||||||| |||||| |||||||||||:::||:||||||| ||||:||||||||||
Db  2408   GGTAGCGCCTTTYTAACCCTCTTGCATTSSTKACGCCTGCTT-AGCTTTGACACCGAGA  2466

Qy  2655   TAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCATCCTCCTATCTTTGGCT  2714
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2467   TAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCATCCTCCTATCTTTGGCT  2526

Qy  2715   TCTTTGCTCACTTGTTACCTCGCTGTGTCGTGTTAGTTAACTCCTATCTTTGGCAACGTTGGG  2774
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2527   TCTTTGCTCACTTGTTACCTCGCTGTGTCGTGTTAGTTAACTCCTATCTTTGGCAACGTTGGG  2586

Qy  2775   AGAATTGGTTTTGGAAGTTACTAAGACCGGAGAGTTTTCCTTGTGCTGGTTGTT  2834
           ||||||||||||||||||||||||||||||||||||| ||||::||||||||:|
Db  2587   AGAATTGGTTTTGGAAGTTACTAAGACCGGAGAGTTTCCTTGYGCTGGTTGTT  2646

Qy  2835   TCCCCGGTGCGACATATGACGCGCTGGTGACTTTCTGTGTGTCAGTAGCTCTCTAT  2894
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2647   TCCCCGGTGCGACATATGACGCGCTGGTGACWTTCTGTGTGTCAGTAGCTCTCTAT  2706

Qy  2895   GTTTAACATCCAGTGCAGCATCGTTCTTGGGACTGACTCGTTAGGGCCCATAGAA  2954
           |||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db  2707   GTTTAACATCCAGTGCAGCAYMGTTCTTGGGACTGACTCGTTAGGGCCCATAGAA  2766

Qy  2955   TGTTGGTGCGTCTCGGAAAGTGTCATGCTTGGTATTCCATTATGTTCTTAAGTTTTTCC  3014
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2767   TGTTGGTGCGTCTCGGAAAGTGTCATGCTTGGTATTCATTATGTTCTTAAGTTTTTCC  2826

Qy  3015   TCTTAGTGTTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCATGGTGATGTCTTGC  3074
           |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db  2827   TCTTAGTGTTTGGTGAGAATGGTGTGTTTTCTAKAAGCACTTGCATGGTGATGTCTTGC  2886

Qy  3075   CTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCATTTTCCCTTTTGAAGGCAAGG  3134
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2887   CTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCATTTTCCCTTTTGAAGGCAAGG  2946

Qy  3135   CAAGGGTCTATAGGAATGAAGGAAGACGCTTGCGCGTGTGGGGACACGGTTGATGGTTGC  3194
           ||||||||||||||||||||||||||||||::::|||||||||||||||||||||||: |
Db  2947   CAAGGGTCTATAGGAATGAAGGAAGACGCTTGGSKKGTGTGGGGACACGGTTGATGGTTTGS  3006

Qy  3195   CCGTTGTTGCGCGTCTCGGCGACACCTTGTTTTCGGCAGGGTTGGCTATGCCGCCAGATGGGT  3254
           |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3007   SCGTTGTTGBGCGCGTCTCGGCGACACCTTGTTTTCCGCAGGGTTAGCTATGCCGCCAGATGGT  3066

Qy  3255   GGGCCAATTACCGCACCTTTACGGCTGGCTGACTGTCTCTGCAGCGTCTGACAACGCGTGTCAGCGA  3314
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3067   GGGCCAATTACCGCACCTTTACGGCTGGCTGACTGTCTCTGCAGCGTCTGACAACGCGTGTCAGCGA  3126

Qy  3315   TGGCAGTGGTCATGATGATAGACACCCCAGAACTTGACTTGAAATCTTCAGATTAG  3374
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3127   TGGCAGTGGTCATGATGATAGACACCCCAGAACTTGACTTGAAATCTTCAGATTAG  3186

Qy  3375   GATCTCTGGCCACTAGCTACATGGGATTTGTTGTGACAACGTGTGTATACTGCTCACC  3434
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3187   GATCTCTGGCCACTAGCTACATGGGATTTGTTGTGACAACGTGTGAATACTGCTCACC  3246

Qy  3435   ATGGCAGCAGCAAGGGGCGGCGGCGGTTGGCTCATCCCAGGCTCATACACCCAATAACGTTG  3494
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3247   ATGGCAGCAGCAACGGGCGGCGGCGGTTGGCTCATCCCACAGCTCCAATAACACCCAATAACCGTTG  3306

Qy  3495   ACGCGGCTAATGACCAGGACATCTATCAACAACATGTGGAGCTGGGTCCCTTACTCGGT  3554
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3307   ACGCGGCTAATGACCAGGACATCTATCAACAACATGTGGAGCTGGGTCCCTTACTCGGT  3366

Qy  3555   GCTCTTGCTCGGGGAGACCAAGGGCCGCTGGGTATCTGTTTTGTGGTTCATTGGTTGAGGTCA  3614
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3367   GCTCTTGCTCGCGGGAGACCAAGGGGTATCTGTTTTGTGGTTCATTGGTTGAGGTCA  3426

Qy  3615   ACAAATCCGATGACCCTTATTGGTGTGTGCGGGGGCCCTTCCCATGGCTGTTGCCAAGG  3674
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db                                                                  .
```

Right column:

```
Db  3427   ACAAATCCGATGACCCTTATTGGTGTGTGCGGGGGCCCTTCCCATGGCTGTTGCCAAGG  3486

Qy  3675   GTTCTTCAGGTGCCCCGATTCTGTGCTCCTCCGGGCATGCATGTTATTGGGATGTTCACCGCTG  3734
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3487   GTTCTTCAGGTGCCCCGATTCTGTGCTCCTCCGGGCATGCATGTTATTGGGATGTTCACCGCTG  3546

Qy  3735   CTAGAAATTCTGGCGGTCAGTCAGTCATTAGTTGTCCGTTGGTGTGTGTGTGCTGGAT  3794
           ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db  3547   CTAGAAATTCTGGCGGTCAGTCAGTCATTAGGGTTAGGCCCGTTGGTGTGTGTGTGCTGGAT  3606

Qy  3795   ACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTACTGTGCCTAACGAGT  3854
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3607   ACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTACTGTGCCTAACGAGT  3666

Qy  3855   ATTCAGTGCAAATTTTAATTGCCCCCACTGGCAGCGGCGAAGTCAACCAAAATTACCACTTT  3914
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3667   ATTCAGTGCAAATTTTAATTGCCCCCACTGGCAGCGGCGAAGTCAACCAAAATTACCACTTT  3726

Qy  3915   CTTACATGAGAGAAGTATGAGGTCTTGGTCTCCAGTCCTAAATCCCAGTGGCTACAACAGCAT  3974
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3727   CTTACATGCAGGRGAAGYATGAGGTCTTGGTCCTAAATCCCAGTGGCTACAACAGCAT  3786

Qy  3975   CAATGCCAAAGTACATGCACGCGACGTACGGCGGTGAATCCAAATTGCTATTTTAATGGCA  4034
           |||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db  3787   CAATGCCAAAGTACATGCACGCGACGTACGGCGGTGAATCCAAATTGCTATTTTAATGGCA  3846

Qy  4035   AATGTACCAACACACAGGGCTTCACTTACGTACGATACAGCACATATGGCATGTGACCGGAG  4094
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3847   AATGTACCAACACACAGGGCTTCACTTACGTACGATACAGCACATATGGCATGTGACCGGAC  3906

Qy  4095   CATGTTCCCGGACTATGATGTAATCATTGTGACGAAGCCATTGTGACGTACCGATGCAACCA  4154
           |||||||||||||||||||||||:|||:|||||||||||||||||||||||||||||||
Db  3907   GATGTTCCCGGAGTAATGATGTAATCATTTGTGACGAATGCGATGTACCGATCGAACCA  3966

Qy  4155   CCGTGTTGGGCATTGGAAAGGTCCTAACCGAAGCTCCATCCAAAAATGTTAGGCTAGTGG  4214
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3967   CCGTGTTGGGCATTGGAAAGGTCCTAACCGAAGCTCCATCCAAAAATGTTAGGCTAGTGG  4026

Qy  4215   TTCTTGCCACGGCTACCCCCCGGAGTAATCCTCACACCACCATGCCAACATAACTGAGA  4274
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4027   TTCTTGCCACGGCTACCCCCCGGAGTAATCCCTCACACCACCATGCCAACATAACTGAGA  4086

Qy  4275   TTCAATTAACCGATGAAGCACTATCCCCTTTCATGGAAAAAGTAAGGAGGAAAATC  4334
           |||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db  4087   TTCAATTAACYGATGAAGCACTATCCCCTTTCATGGAAAAAGTAAGGAGGAAAATC  4146

Qy  4335   TGAAGAAAGGGGAGACACCTTATCTTTGAGGCTACCAAAAAACACTGTGATGAGCTGCTA  4394
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4147   TGAAGAAAGGGAGACACCTTATCTTTGAGGCTACCAAAAAACACTGTGATGAGCTTGCTA  4206

Qy  4395   ACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTAGGGGATGTGACATCTCAA  4454
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4207   ACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTAGGGGATGTGACATCTCAA  4266

Qy  4455   AAATCCCTGAGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTGTCCACTGTACACTG  4514
           ||||::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4267   AAATGCCTGAGGGCGCCGCACAGGGCTGGGGGGAGAGCTGGCATATACTACTTACACTG  4326

Qy  4515   GTGACTTTGATTCCGTGTATGACTGCAGCCTCATGGTGAAGGCACATGCCATGTTGACC  4574
           |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4327   GTGACTTTGATTCCGTGTATGACTGCAGCCTCATGGTGAAGGCACATGCCATGTTGACC  4386

Qy  4575   TTGACCCTACTTTCACCATGTGGTGTTCGTGTGTGTGCAATAGTTAAAGGCC  4634
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4387   TTGACCCTACTTTCACCATGTGGTGTTCGTGTGTGTGCAATAGTTAAAGGCC  4446

Qy  4635   AGCGTAGGGGCCGCCACAGCCGTGGGAGAGCTGGCATATACTACTATGTAGACGGGAGTT  4694
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4447   AGCGTAGGGGCGCCCACACAGCCGTGGGAGAGCTGGCATATACTACTATGTAGACGGGAGTT  4506

Qy  4695   GTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTTCGACGCAGCCAAGG  4754
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4507   GTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTTCGACGCAGCCAAGG  4566
```

Left column:

```
Qy  4755  CATGGTATGGTTTGTCATCAACAGAGAAGCTCAAACTATTCTGGACACCATCGCACCCAAC  4814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4567  CATGGTATGGTTTGTCATCAACAGAGAAGCTCAAACTATTCTGGACACCATCGCACCCAAC  4626

Qy  4815  CTGGGTTACCTGCGATAGGAGCAAATTGGACGAGTGGCTGATCTCTTTTCTATGGTCA  4874
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4627  CTGGGTTACCTGCGATAGGAGCAAATTGGACGAGTGGCTGATCTCTTTTCTATGGTCA  4686

Qy  4875  ACCCCGAACCTTCATTGTCAATACTGCGACAAAAGAACTGCTGACAATTATGTTTTGTTGA  4934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4687  ACCCCGAACCTTCATTGTCAATACTGCGACAAAAGAACTGCTGACAATTATGTTTTGTTGA  4746

Qy  4935  CTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCCCAATGACGCACCAC  4994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4747  CTGCAGCCCAACTACAACTGTGTCATCAGTATGGCTATGCTGCTCCCAATGACGCACCAC  4806

Qy  4995  GGTGGCAGGGAGCCCGCTTGGGAAAAAACCTTGTGTGGCGTTCGTGGCGCTTGGACGGCG  5054
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4807  GGTGGCAGGGAGCCCGCTTGGGAAAAAACCTTGTGTGGCGTTCGTGGCGCTTGGACGGCT  4866

Qy  5055  CTGACGGCCTGTCCTGGCGCCAGAGCCAGGAGGTGACCAGATACCAAATGTGCTTCACTG  5114
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4867  GTGACGGCCTGTCCTGGCGCCAGAGCCAGGAGGTGACCAGATACCAAATGTGCTTCACTG  4926

Qy  5115  AAGTCAATACTTCTGGGACAGCCGCACTCGCTGTTGCGCTGGCTATGGCTTATC  5174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4927  AAGTCAATACTTCTGGGACAGCCGCACTCGCTGTTGCGCTGGCTATGGCTTATC  4986

Qy  5175  TAGCCATTGACACTTTTGGCGCCACTTGTGTTGCGTCGGTCTATTACATCAGTCC  5234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4987  TAGCCATTGACACTTTTGGCGCCACTTGTGTTGCGTCGGTCTATTACATCAGTCC  5046

Qy  5235  CTACCGGTGCTACTGTCGCGCCCCAGTGGTTGACGAAGAAGAAATGTGGAGGAGTGTGCAT  5294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5047  CTACCGGTGCTACTGTCGCGCCCCAGTGGTTGACGAAGAAGAAATGTGGAGGAGTGTGCAT  5106

Qy  5295  CATTCATTCCCTTGGAGGCCATGGTTGCTGCAATTGACAAGAGTACAATCACCA  5354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5107  CATTCATTCCCTTGGAGGCCATGGTTGCTGCAATTGACAAGAGTACAATCACCA  5166

Qy  5355  CAACTAGTCCTTCTTCACATTGGAAACGCCCCTGAAAAACTTAACACCTTTCTTGGCCTC  5414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5167  CAACTAGTCCTTCTTCACATTGGAAACGCCCCTGAAAAACTTAACACCTTTCTTGGCCTC  5226

Qy  5415  ATGCAGCTACAAATCCTTGCTATCAGAGTATTGCGGGTATTACACCCCACTACCTCACA  5474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5227  ATGCAGCTACAAATCCTTGCTATCAGAGTATTGCGGGTATTACACCCCACTACCTCACA  5286

Qy  5475  ATCCCTTTGCATCATGCGTGTTTGCTTCATTGCGGGCTATGCGCCACTATGCTGCTTCT  5534
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5287  ATCCCTTTGCATCATGCGTGTTTGCTTCATTGCGGGCTATGCGCCACTTACCTGCTTCT  5346

Qy  5535  AGATCAAAATGTTCCTGTCATTATTTGGAGGGCAATTGCGTCCAAGCTTACAGACGCTA  5594
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5347  AGATCAAAATGTTCCTGTCATTATTTGGAGGGCAATTGCGTCCAAGCTTACAGACGCTA  5406

Qy  5595  GAGGGCACTGGCGTTCATGATGGCCCGGGGGCTGCGGGAACAGCTCTTGGTACATGGACAT  5654
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5407  GAGCGGCACTGGCGTTCATGATGGCCCGGGGGCTGCYGGGAACAGCTCTTGGTACATGGACAT  5466

Qy  5655  CGGTGGGTTTTGTCTTTGCGTGTTGCTTCATTGCCGGCTATGCTGCCGCCCTCATCCACTGCTTGCT  5714
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5467  CGGTGGGTTTTGTCTTTGCGTGTTGCTTCATTGCCGGCTATGCTGCCGCCCTCATCCACTGCTTGCT  5526

Qy  5715  TGACATTTAAATGCTTGATGGGTGAGTGGGTGGCCCACTATGGACTGCTGTTCGTTAGTTCT  5774
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5527  TGACATTTAAATGCTTGATGGGTGAGTGGGTGGGCYCACTATGGACTGCTGTTCGTTAGTTCT  5586

Qy  5775  ACTCCGCGTTCAATCCGGCCGCCAGGAGTTGTGGGCGTTGGGCGTTGTGGCCGTTGTGCAATGTTTG  5834
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5587  ACTCCGCGTTCAATCCGGCCGCCAGGAGTTGTGGGCGTTGGGCGTTGTGGCCGTTGTGCAATGTTTG  5646
```

Right column:

```
Qy  5835  CTTTGACAACGCAGCAGGGCCAGATCACTGGCGCCCAACAGACTTCTTACTATGCTGCTAGGA  5894
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5647  CTTTGACAACGCAGCAGGGCCAGATCACTGGCGCCCAACAGACTTCTTACTATGCTGCTAGGA  5706

Qy  5895  GCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCGAGCAGAAGATACTGG  5954
          |||||||||||||::|||||||||||||||||||||||||||||||||||||||||||||
Db  5707  GCAACACTGTATGTARTGAGTACTTTATTGCCACTCGTGACATCCGAGCAGAAGATACTGG  5766

Qy  5955  GCATTCTTGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCTCATCCGTTGCTTCCACA  6014
          |||||||||||||||||||||||||||||||||::|||||||||||||:|||::|:|||||
Db  5767  GCATTCTTGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCTCATCCGTTGCTYCACA  5826

Qy  6015  CCCCGACGGCGGAGATGATTGCGGCCTCATTGCTTGGGGTTCTACAGATTTGGCAGTATGTGT  6074
          ||||||||||||||||||||||||||||||||||||||||||||||::|||:|:|||||||
Db  5827  CCCCGACGGCGGAGATGATTGCGGCCTCATTGCTTGGGGTCTAARAGATTTGGCAGTATGTGT  5886

Qy  6075  GCAATTTCTTGTGATTGCTTAATGTCCTTAAAGTGGAGTTCAGAGCACGTGTTAACA  6134
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5887  GCAATTTCTTGTGATTGCTTAATGTCCTTAAAGTGGAGTTCAGAGCACGTGTTAACA  5946

Qy  6135  TTCCTGGTTGTCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCCCTGGATTGGATCAG  6194
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5947  TTCCTGGTTGTCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCCCTGGATTGGATCAG  6006

Qy  6195  GTATGTCCAAGCACGCGTGTCCATGCGGCGTGAACTCATCTTTCTGTGTTGAGAATGGTT  6254
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6007  GTATGTCCAAGCACGCGTGTCCATGCGGCGTGAACTCATCTTTCTGTGTTGAGAATGGTT  6066

Qy  6255  TTGCAAAACTTACAAAGGACCCCAGAACTGTTCAAAATACTGGAGAGGGGCGCTGTTCCAG  6314
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6067  TTGCAAAACTTACAAAGGACCCCAGAACTGTTCAAAATACTGGAGAGGGGCGCTGTTCCAG  6126

Qy  6315  TCAACGCTAGGCGTGTGGGTCGGCTAGACCGACCCAACTGTTCAAATTACTGATTGGATCTCG  6374
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6127  TCAACGCTAGGCGTGTGGGTCGGCTAGACCGACCCAACTGTTCAAATTACTGATTGGATCTCG  6186

Qy  6375  TCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGATCACATTTTTGTA  6434
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6187  TCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGATCACATTTTTGTTA  6246

Qy  6435  CAGCAGTATCCTCCTCCAAATGTCTGTTCACCCAGTTGCCCCCAACCTTGAGAGCTGCAG  6494
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6247  CAGCAGTATCCTCCTCCAAATGTCGTTCACCCAGTTGCCCCCAACCTTGAGAGCTGCAG  6306

Qy  6495  TGGCCGTGGACGGCGTACAGGTTCAGTTGTTATCTGGAGCGTAAGCGTAAAACTCCTTGGACGA  6554
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6307  TGGCCGTGGACGGCGTACAGGTTCAGYGTATCTGGAGCGTAAGCGTAAAACTCCTTCCCCCG  6366

Qy  6555  CATCTGCTTGCTGTCGTGTTTCATTGCGCCGGTATTACACCCCACTACCTCACA  6614
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6367  CATCTGCTTGCTGTCGTGTTTACGGTCTGCGACGGTAAGCTGTTAAGCTTCCCTTCCCCG  6426

Qy  6615  TTGACGGTCACACACCTGGTGTGCGCATGCAACTTAATTTGCCTGATGCACTTGAGACAA  6674
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6427  TTGACGGACACACACCTGGTGTGCGCATGCAACTTAATTTGCCTGATGCACTTGAGGCAA  6486

Qy  6675  ATGACTGTAATTCCACAAACAATCCTCCAGTGATGAGACCGGAGTGTCCGGCTCTTGTTT  6734
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6487  ATGACTGTAATTCCATAAACAACACTCCTAGTGATGAGAAGCCCAGTGTCCGGCTCTTGTTT  6546

Qy  6735  TCAAACAGGAGTTGCGGCGTACAAAACCAATTGCTTGAGGCAATTTCAGCTGGCGTTGACA  6794
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6547  TCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTCAGCTGGCGTTGACA  6606

Qy  6795  CACCCAAACTGCCAGCCCCCTCC---ATCGAAGAGAGTAGTGTAAGAAAGCCCAGTTCC  6851
          |||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db  6607  CCACCAAACTGCCAGCCCCCTCCCCAGATCGAAGATCGAAGATCGAAGGTAGTGTAAGAAAGCCCAGTTCC  6666

Qy  6852  GGGCAAGAACTGGTTCGGCTTCGGGGGTTGTGGGCGTCTTCGCTCCCCCCCTCCG  6911
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6667  GGGCAAGAACTGGTTCGCTTGCCTTGCCTCCCTCCGAGATCGTCCCCAGGAGTGTCAT  6726

Qy  6912  GTCCTGAAAGCCTGCAACGAAGTGACCCGGTTAGAAGTCTTCAAACCTCCCTCCTTCAC  6971
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       GGGCAAGAACTGGTTCGCTTGCCTCCCTCCGAGATCGTCCCCAGGAGTGTCAT  6726
```

```
Db  6727 ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||| 6786
         GTCCTGAAAGCCTGCAACGAAGTGACCCGTTAGAAGGTCCTTCAAMCCTCCCTTCTTCAC
Qy  8052 CGATGATGCGCTTATGATGCGGCCGAGAGATCGGATCGGATACGCTGTAGGTCTTCCGGCGTCT 8111

Db  6972 CACCTGTTCTACAGTTGGCCATGGCCGATGCCCGTTGGGAGCGGGGTGAGTGTAACCCTT 7031
         ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Qy       CACCTGTTCTRCAGTTGGCCATGGCCGATGCCCGTTGGGAGCAGGGTGAGTGTAACCCTT 6846

Db  7032 TCACTGCAATTGGATGTGTCAATGACCGAAACAGGCGGAGGCCCTGATGATTACCCAGTT 7091
         |||||||||||||||||||||||||||||||| |:||||||||||MKRAATTACCCAGTT
Qy       TCACTGCAATTGGATGTGTCAATGACCGAAACARGYGGAGKCCCWMAKRAATTACCCAGTT 7986

Db  7092 ACCCTCCCAAAAAGGAGGAGTCTCTGAATGGTCGCAGCAGAAAGTTGGTCGACGGCTACAACCG 7151
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       ACCCTCCCAAAAAGGAGGAGTCTCTGAATGGTCGCAGCAGAAAGTTGGTCAACGACTACAACCG 6906

Db  7152 CTTCCAGTACGTTACTGGCCCCCCGTACCCTAAGATACGGGAAAGATTCCACTCAGT 7211
         ||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Qy       CTTCCAGTACGTTCRCAGTTACTGGCCCCCCGTACCCTAAGATACGGGAAAGGATTCCACTCAAT 7026

Db  7212 CAGCCCCCGCCAAACGGCCTACAAAAAGAAGTTGGGAAAGAAGTGAGTTTCGTGCAGCA 7271
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       CAGCCACCGCCAAACGGCCTACAAAAAGAAGTTGGGAAAGAAGTGAGTTTTCGTGCAGCA 7086

Db  7272 TGAGTACTACACCTGGACAGCGTGATTAGCTTCAAAACTGCTTCAAAACTGCTTCTGTCTGCAA 7331
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       TGAGTACTACACCTGGACCACGTGATTAGCTTCAAAACTGCTTCAAAACTGCTTCTGTCTGCAA 7146

Db  7332 CTCGGGGCCATCACTAGTTGGTTCCTCAAACAAGATCATTGGTATGTGACTGAGCCGC 7391
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       CTCGGGGCCATCACTAGTTGGTTCCTCAAACAAGATCATTGGTATGTGACTGAGCCGC 7206

Db  7392 GGGATGCGGAGCTTAGAAAACAAAAGTCACTATTAATAGACAACCTCTGTTCCCCCAT 7451
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       GGGATGCGGAGCTTAGAAAACAAAAGTCACTATTAATAGACAACCTCTGTTCCCCCAT 7266

Db  7452 CATACCACAAGCAAGTGAGAGTTGGCTAAGGAGGAAAAAGCTTCAAAAGCTTCAAAAGG 7511
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       CATACCACAAGCAAGTGAGAGTTGGCTAAGGAGGAAAAAGCTTCAAAAGCTTCAAAAGG 7326

Db  7512 GGGACTATGATGAAGTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTCTCCACATCACTG 7571
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       GGGACTATGATGAAGTAGCAGCTCACACGCCCTCTAAGTCTGCTAAGTCTCCACATCACTG 7386

Db  7572 GCCTTCGGGGCACTGATGTTCGTTCGTTCTGGACTTGCAGA 7631
         ||||||||||||||||||||||||||||||||||||||||||
Qy       GCCTTCGGGGCACTGATGTTCGTTCGTTCTGGACTTGCAGA 7446

Db  7632 AGTGTGTCGAGCGAGTTCATGTCGAGTACCGAGTCATTATCGGCAAACTGTGATAGTTCCAAAGG 7691
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       AGTGTGTCGAGCGAGTTCATGTCGAGTACCGAGTCATTATCGGCAAACTGTGATAGTTCCAAAGG 7506

Db  7692 AGGAGGTCTTCGTGAAGACCCCCAGAAACAACAAGAAACCCCCAGAAACAACAAGAAACCCCCAGA 7751
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       AGGAGGTCTTCGTGAAGACCCCCAGAAACAACAAGAAACCCCCAGAAACAACAAGAAACCCCCAGA 7566

Db  7752 ACCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGTTGCTCCTGACG 7811
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       ACCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGTTGCTCCTGACG 7626

Db  7812 TAGTTAAAGCTGTCGAGCGAGGTCGTACGGGTTTGTTGATAGTTCCAAAGG 7871
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       TAGTTAAAGCTGTCGAGCGAGGTCGTACGGGTTTGTTGATCCGCAAACTGTGTCAAGC 7686

Db  7872 GTCTGTGTCGATGTGGTCACCCGAGCGCCACATGCATCAGTGTGTTTG 7931
         ||||||||||||||||||||||||||||||||||||||||||||||||
Qy       GTCTGTGTCGATGTGGTCACCCGAGCGCCACATGCATCAGTGTGTTTG 7746

Db  7932 ACAGTACCATCACACCCGAGATATCATCGTGGAGACAGAGTATCTACTCACTGCAGCTAAAC 7991
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       ACAGTACCATCACACCCGAGATATCATCGTGGAGACAGAGTATCTACTCACTGCAGCTAAAC 7806

Db  7992 TCAGTGACCAACACCGAGCTGGCATTCACCACCATTGCGAGGCAGTATCACGCTGGAGGAC 8051
         ||||||||||||||||||||||||||||||||||||
Qy       TCAGTGACCAACACCGAGCTGGCATTCACCACCATTGCGAGGCAGTATCACGCTGGAGGAC 8051
```

```
Db  7807 TCAGTGACCAACACCGAGCTGGCATTCACCACCATTGCGAGGCACTATCACGCTGGAGGAC 7866
         |||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Qy  8052 CGATGATGCGCTTATGATGCGGCCGAGAGATCGGATCGGATACGCTGTAGGTCTTCCGGCGTCT 8111

Db  7867 CGATGATGCGCTTATGATGCGGCCGAGAGATCGGATCGGATACGCTGTAGGTCTTCCGGCGTCT 7926
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       CACCTGTTCTRCAGTTGGCCGATGCCCGTTGGGAGCAGGGTGAGTGTAACCCTT 6846

Db  8112 ATACTACCTCAAGTCCAACAGTTGACCTGCTGGCTGGCTGAAGTAAATGCTGCAGCCGAAC 8171
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  7927 ATACTACCTCAAGTCCAACAGTTGACCTGCTGGCTGGCTGAAGTAAATGCTGCAGCCGAAC 7986

Db  8172 AGGCTGGCATGGAAGAACCCTCGCTTCCTTATTTGCGGGCGATGATTGCACCGTAATTTGGA 8231
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  7987 AGGCTGGCATGGAAGAACCCTCGCTTCCTTATTTGCGGGCGATGATTGCACCGTAATTTGGA 8046

Db  8232 AGAGCGCCGGAGCGAGCAGATGCAGACGAAGCAATGGGTGTCTTTGCTAGCTGGATGATGAAGG 8291
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  8047 AGAGCGCCGGAGCGAGCAGATGCAGACGAAGCAATGGGTGTCTTTGCTAGCTGGATGATGAAGG 8106

Db  8292 TGATGGGTGCACCACAGATTGTGTGCCTCAACCCAAATACAGTTGGAAGAATTAACAT 8351
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  8107 TGATGGGTGCACCACAGATTGTGTGCCTCAACCCAAATACAGTTGGAAGAATTAACAT 8166

Db  8352 CATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGGCCTTACTACTTTCTTA 8411
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  8167 CATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTACTACTTTCTTA 8226

Db  8412 CAAGAGATCCTGCTCCGTATCCCCCTGGCAGGTGCTCTGCCGCAGGGTCTGGGATACAACCCCA 8471
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  8227 CAAGAGATCCTGCTCCGTATCCCCCTGGCAGGTGCTCTGCCGCAGGGTCTGGGATACAACCCCA 8286

Db  8472 GTGTGCGTGGATTGGGTATCATCAATAACATCATCACTACCCATGTTGTGGGGTTAGCCGTGT 8531
         |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  8287 GKGCKGCGTGGATTGGGTATCATCAATAACATCATCACTACCCATGTTGTGGGGTTAGCCGTGT 8346

Db  8532 TGGCTGTCCATTTCATGGTCGAGCAGCAGATGCTCTCTTGAGGACAAACTTCCCGAGACTGTGACCT 8591
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  8347 TGGCTGTCCATTTCATGGTCGAGCAGCAGACAGATGCTCTCTTTGAGGACAAACTTCCCGAGACTGTGACCT 8406

Db  8592 TTGACTGGTATGGGAAAAATTATACGGCCTGTAGAGAATCGCTG 8651
         ||||||||||||||||||||||||||||||||||||||||||||
Qy  8407 TTGACTGGTATGGGAAAAATTATACGGCCTGTAGAGAATCGCTG 8466

Db  8652 GTGTGCACGGTATTGAGGCTTTCTCGGTGGTCGTGCCTACACCAACGCTGCAGATCCTCAGAG 8711
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  8467 GTGTGCACGGTATTGAGGCTTTCTCGGTGGTCGTGCCTACACCAACGCTGCAGATCCTCAGAG 8526

Db  8712 TTTCCCAATCACTAACAGACATGACCATGCCCCCCCTGCGAGCCTGGGCGAAAGAAAGCCA 8771
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  8527 TTTCCCAATCACTAACAGACATGACCATGCCCCCCCTGCGAGCCTGGGCGAAAGAAAGCCA 8586

Db  8772 GGGGCGGTCCTGCCGCCCAAGAGGCGTGGCGGAGCACACGCACAG-AAAATTGGCTGCTTCC 8831
         |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Qy  8587 GGGCGGTCCTGCCGCCCAAGAGGCGTGGCGGAGCACACG-AAAATTGGCTGCTTCC 8645

Db  8832 TTCTCTGGCATGCTACATCTAGAACTGGATGTTTACTCCCCGGACTATAAGACGAGCGTGGCTCGGT 8891
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  8646 TTCTCTGGCATGCTACATCTAGAACTGGATGTTTACTCCCCGGACTATAAGACGAGCGTGGCTCGGT 8705

Db  8892 ACACCACTTTCAATTATTGTGATGTTTACT-CCCSGACGRGGACGTGTTTATTACACCAC 8951
         ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Qy  8706 ACACCACTTTCAATTATTGTGATGTTTACT-CCCSGACGRGGACGTGTTTATTACACCAC 8764

Db  8952 AGAGAAGAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTTGCCCTAGGGC 9011
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  8765 AGAGAAGAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTTGTGTGCCCTAGGGC 8824

Db  9012 TCATTGCTGTTGGATTAGCCATCAGCTGGAACAGACATCTACTCACGCAAATTCAAAATTAACTAACAGTTTT 9071
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  8825 TCATTGCTGTTGGACTAGCCATCAGCTGGACAGACATCTACTCACGCAAATTCAAAATTAACTAACAGTTTT 8884

Db  9072 TTTTTTTTTTTTTTTTTTT 9090
         |||||||||
Qy  8885 TTTTTTTTTTTTTTTTTTTTTTT 8903
```

RESULT 14
US-08-469-260A-80
; Sequence 80, Application US/08469260A
; Patent No. 6451578
GENERAL INFORMATION:
;   APPLICANT: JOHN N. SIMONS
;   APPLICANT: TAMI J. PILOT-MATIAS
;   APPLICANT: GEORGE J. DAWSON
;   APPLICANT: GEORGE G. SCHLAUDER
;   APPLICANT: SURESH M. DESAI
;   APPLICANT: THOMAS P. LEARY
;   APPLICANT: ANTHONY SCOTT MUERHOFF
;   APPLICANT: JAMES C. ERKER
;   APPLICANT: SHERI L. BUIJK
;   APPLICANT: ISA K. MUSHAHWAR
;   TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
;   TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
;   NUMBER OF SEQUENCES: 716
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
;   STREET: 100 ABBOTT PARK ROAD
;   CITY: ABBOTT PARK
;   STATE: IL
;   COUNTRY: USA
;   ZIP: 60064-3500
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/469,260A
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/08/424,550
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: POREMBSKI, PRISCILLA E.
;   REGISTRATION NUMBER: 33,207
;   REFERENCE/DOCKET NUMBER: 5527.PC.01
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 708-937-6365
;   TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 80:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4268 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-08-469-260A-80

Query Match   44.2%; Score 4159; DB 3; Length 4268;
Best Local Similarity   98.1%; Pred. No. 0;
Matches 4215; Conservative 20; Mismatches 30; Indels 30; Gaps 2;

Db   181 GGTGTGTGTGCGGGGCCCTTCCATGCTGCGTGTGCCAAGGGTTCTTCAGGTGCCCGATTC 240

Qy 3696 TGTGCTCCTCCGGGCATGTTATTGGGATGGTTCACCGCTGCTAGAAATTCTGGCGGTTCAG 3755
         || |||||| |||||||||||||||||||||||||| ||||||||||||||||||||||
Db   241 TGTGCTCCTCCGGGCATGTTATTGGGATGGTTCACCGCTGCTAGAAATTCTGGCGGTTCAG 300

Qy 3756 TCAGTCAGATTAGGGTTAGGCGTTGTGTGTGTCTGATACCATCCCAGTACACAGCAC 3815
         || ||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TCGGCCCAGATTAGGGTTAGGCGTTGTGTGTGTCTGATACCATCCCAGTACACAGCAC 360

Qy 3816 ATGCCACTCTTGATACAAAACCTACTGTGCCTAACGAGTATTCAGTGCAAATTTAATTG 3875
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 ATGCCACTCTTGATACAAAACCTACTGTGCCTAACGAGTATTCAGTGCAAATTTAATTG 420

Qy 3876 CCCCCACTGGCACGCGGCAAGTCAACCAAATTACCACTTTCTTACATGCAGGAGGAAGTATG 3935
         |||||||||||||||||||||||||||||||||||||||||||||||||: |||||:|||
Db   421 CCCCCACTGGCACGCGGCAAGTCAACCAAATTACCACTTTCTTACATGCAGGCGGAAGYATG 480

Qy 3936 AGGTCTTGGTCCTAAATCCCAGTGTGGCTACAACAGCATCAATGCCAAAGTACATGCACG 3995
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AGGTCTTGGTCCTAAATCCCAGTGTGGCTACAACAGCATCAATGCCAAAGTACATGCACG 540

Qy 3996 CGAGCGTACGGCGTGAATCCAAATGCTATTTAAATGGCAAATTGTACCACACAGGGGCTT 4055
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CGAGCGTACGGCGTGAATCCAAATGCTATTTAAATGGCAAATTGTACCACACAGGGGCTT 600

Qy 4056 CACTTAGTACAGACACATATGGCATGTACCTGACCGGAGCATGTTCCCGGAACTATGATG 4115
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CACTTAGTACAGACACATATGGCATGTACCTGACCGGAGCGATGTTCCCGGAACTATGATG 660

Qy 4116 TAATCATTGTGACGAATGCCATGCTACCACCGTGTCAACCACCGTGTTGGGAAGG 4175
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TAATCATTGTGACGAATGCCATGCTACCACCGTGTCAACCACCGTGTTGGGAAGG 720

Qy 4176 TCCTAACCGAAGCTCCATCCAAAAATGTTAGGCTAGTTCTTGCCACGGCTACCCCCC 4235
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TCCTAACCGAAGCTCCATCCAAAAATGTTAGGCTAGTTCTTGCCACGGCTACCCCCC 780

Qy 4236 CTGGAGTAATCCCTACCACATACCAACATACTGAGATTCAATTAACCGATGAAGGCA 4295
         |||||||||||||||||||||||||||||||||||||||||:||||||:|||||||||
Db   781 CTGGAGTAATCCCTACCACATACCAACATACTGAGATTCAATTAACYGATGAAGGCA 840

Qy 4296 CTATCCCCTTCATGGAAAAAGATTAAGGAGGAAAATCTGAAGAAAGGAGACACCTTA 4355
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CTATCCCCTTCATGGAAAAAGATTAAGGAGGAAAATCTGAAGAAAGGAGACACCTTA 900

Qy 4356 TCTTTGAGGCTACCAAAAAACACTGTGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAA 4415
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCTTTGAGGCTACCAAAAAACACTGTGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAA 960

Qy 4416 TAACAGCTGTCTCTTACTATAGGGGATGTGACATCTCAAAAATCCCTGAGGGGACTGTG 4475
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TAACAGCTGTCTCTTACTATAGGGGATGTGACATCTCAAAAATGCCTGAGGGGACTGTG 1020

Qy 4476 TAGTAGTTGCCACTGATGCCTTGTGTACAGGTACACTGGTACTTGATTCCGTGTATG 4535
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 TAGTAGTTGCCACTGATGCCTTGTGTACAGGTACACTGGTACTTGATTCCGTGTATG 1080

Qy 4536 ACTGCAGCCTCATGGTAGAAGGCACATGCCATGTTGACCTTGACCCTACTTTCACCATGG 4595
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 ACTGCAGCCTCATGGTAGAAGGCACATGCCATGTTGACCTTGACCCTACTTTCACCATGG 1140

Qy 4596 GTGTTCGTGTGTGCGGGGTTTCAGCAATAGTTAAAGGCCCAGCTAGGGGCCCACAGG 4655
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 GTGTTCGTGTGTGCGGGGTTTCAGCAATAGTTAAAGGCCCAGCTAGGGGCCCACAGG 1200

Qy 4656 GTGGGAGAGCTGGCATATACTACTATGTGTACCCCTGGAGTTGTACCCCTGGTATGGTTC 4715
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 GTGGGAGAGCTGGCATATACTACTATGTGTACCCCTGGAGTTGTACCCCTGGTATGGTTC 1260

Qy 4716 CTGAATGCAACATTGTTGAAGCCTTCGACGCAGCCAAGGCATGGTATGTTTGTCATCAA 4775
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 CTGAATGCAACATTGTTGAAGCCTTCGACGCAGCCAAGGCATGGTATGTTTGTCATCAA 1320

Qy 3456 TGGCTCATCCCACAGGCTCTATACACCCCAATAACCGTTGACGCGGCTAATGACCAGGACA 3515
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TGGCTCATCCCACAGGCTCTATACACCCCAATAACCGTTGACGCGGCTAATGACCAGGACA 60

Qy 3516 TCTATCAACCACCATGTGGAGCTGGGTCCCTTACTCGGTGCTCTTGCGGGGAGACCCAAGG 3575
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TCTATCAACCACCATGTGGAGCTGGGTCCCTTACTCGGTGCTCTTGCGGGGAGACCCAAGG 120

Qy 3576 GGTATCTGGTAACACGACTGGGGTCATTGGTTGAGGTCAACAAATCCGATGACCCTTATT 3635
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGTATCTGGTAACACGACTGGGGTCATTGGTTGAGGTCAACAAATCCGATGACCCTTATT 180

Qy 3636 GGTGTGTGTGCGGGGCCCTTCCCATGCTGCGTGTGCCAAGGGTTCTTCAGGTGCCCCGATTC 3695
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

us-10-009-002-1.rni

```
Qy  5856 ATCACTGGCCAACAGACTTCTTACTATGCTTGCTAGGAGCAACACTGTATGTAATGAGT 5915
         |||||||||||||||||||||||||||||||||||||||||||||||||: ||||:||||
Db  2401 ATCACTGGCCAACAGACTTCTTACTATGCTTGCTAGGAGCAACACTGTATGTARTGAGT 2460

Qy  5916 ACTTTATTGCCACTCGTGACATCCGCAGGAAGATACTGGGCATTCTGAGGCATCTACCC 5975
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 ACTTTATTGCCACTCGTGACATCCGCAGGAAGATACTGGGCATTCTGGAGGCATCTACCC 2520

Qy  5976 CCTGGAGTGTCATATCAGCTTGCATCCGTGGCTTCACACCCGACGGAGGATGATTGCG 6035
         |||||||: |:||||||||||||||||||||||||: |||||||||||||||||||||
Db  2521 CCTGGAGTRTCATATCAGCTTGCATCCGTGGCTYCACACCCCGACGGAGGATGATTGCG 2580

Qy  6036 GCCTCATTGCTTGGGGTCTTAGAGATTTGGCAGTATGTGCAATTTCTTGTGATTTGCT 6095
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581 GCCTCATTGCTTGGGGTCTARAGATTTGGCAGTATGTGCAATTTCTTGTGATTTGCT 2640

Qy  6096 TTAATGTCCTTAAAGCTGAAGCTGAGTTCAGAGCATGGTAACATTCCTGGTTGTCCTTTCTACA 6155
         ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db  2641 TTAATGTCCTTAAAGCTGAAGCTGAGTTCAGAGCATGGTAACATTCCTGGTTGTCCTTTCTACA 2700

Qy  6156 GCTGCCAGAAGAGGGGTACAAGGGGTACAAGGGGTACAAGCTGATCAGGTAATGCTCCAAGCACGCTGTC 6215
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 GCTGCCAGAAGAGGGGTACAAGGGGTACAAGGGGTACAAGCTGATCAGGTAATGCTCCAAGCACGCTGTC 2760

Qy  6216 CATGCGGTGCTGAACTCATCTTTCTGTTGAGAATGGTTTGCAAAACTTTACAAAGGAC 6275
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 CATGCGGTGCTGAACTCATCTTTCTGTTGAGAATGGTTTGCAAAACTTTACAAAGGAC 2820

Qy  6276 CCAGAACTTGTTCAAATTACTGGAGAGGGCTGTTCAGTCAACGCTAGGCTGTGTGGGGT 6335
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821 CCAGAACTTGTTCAAATTACTGGAGAGGGCTGTTCAGTCAACGCTAGGCTGTGTGGGGT 2880

Qy  6336 CGGCTAGACCGGACCCAACTGATTGGACTAGTCCTGTCGTCAATTATGGCGTTAGGGACT 6395
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881 CGGCTAGACCGGACCCAACTGATTGGACTAGTCCTGTCGTCAATTATGGCGTTAGGGACT 2940

Qy  6396 ACTGTAAATATGAGAAAATGGGAGATCACATTTTTGTACAGCAGTATCCTCTCCAAATG 6455
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 ACTGTAAATATGAGAAAATGGGAGATCACATTTTTGTACAGCAGTATCCTCTCCAAATG 3000

Qy  6456 TCTGTTTCACCCAGGTGCCCCAACCTTGAGACTGGGCCGTGGGACGGGCCGTGACGGCTACAGG 6515
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 TCTGTTTCACCCAGGTGCCCCAACCTTGAGACTGGGCCGTGGGACGGGCCGTGACGGCTACAGG 3060

Qy  6516 TTCAGTGTTATCTAGGTGAGCCCAAAACTGTTAAGCTTGTCCCTTCCCTTGCGTTGCTTGCTACGGTC 6575
         |||::|||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  3061 TTCAGYGTTATCTAGGTGAGCCCAAAACTGTTAAGCTTGTCCCTTCCCTTGCGTTGCTTGCTACGGTC 3120

Qy  6576 CTGACGGTAAGGGTAAAACTGTTAAGCTTGTTCCTGTGACGGTGACGGACACACCTGGTG 6635
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 CTGACGGTAAGGGTAAAACTGTTAAGCTTGTTCCTGTGACGGTGACGGACACACCTGGTG 3180

Qy  6636 TGCGCATGCAACTTAATTGCGTGATGCACTTGAGACAAAATGACTGTAATTCCACAAACA 6695
         ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db  3181 GTCCGCATGCAACTTAATTGCGTGATGCGATCGACTTGAGGCAAATGACTGTAATTCCATAAAACA 3240

Qy  6696 ACACTCCTAGTGATGAAGCCGCAGTGTCCGCAGTTTCAAACAGGAGTTGCGGCGTA 6755
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 ACACTCCTAGTGATGAAGCCGCAGTGTCCGCTCTTGTTTTCAAACAGGAGTTGCGGCGTA 3300

Qy  6756 CAAACCAATTGCTTGAGGCAATTTCAGCTGGCAATTTTAAATGCTTGATGG 6815
         ||||: |||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 CAAACCAATTGCTTGAGGCAATTTCAGCTGGCAATTTTAAATGCTTGATGG 3360

Qy  6816 CC---ATCGAAGAGGTAGTTGCGTGATGCACTTGAGACACTTGAGAGAAAGCCGCCAGTTCGGTTCGCTTA 6872
         ||   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 CCCAGATCGAAGAGGTAGTTGCGTGATGCACTTGAGACACTTGAGAGAAAGCCGCCAGTTCGGTTCGCTTA 3420

Qy  6873 CCTTGCCTCCCCCCCGAGATCCGTCCCAGGAGTGTCATGTCCTGAAAGCCTGCAACGAA 6932
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421 CCTTGCCTCCCCCCCCGAGATCCGTCCCAGGAGTGTCATGTCCTGAAAGCCTGCAACGAA 3480

Qy  6933 GTGACCCGTTAGAAGGTCCTTCAAACCTCCTTCCTCCAACCTGTTCTACCACCGTTGGCCA 6992
```

```
Qy  4776 CAGAAGCTCAAAACTAATTCTGGACACCTATCGCACCCAACCTGGGTTACCTGCGATAGGAG 4835
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CAGAAGCTCAAAACTAATTCTGGACACCTATCGCACCCAACCTGGGTTACCTGCGATARGAG 1380

Qy  4836 CAAATTTGGACGAGTGGCTGATCTCTTTCTCTTTTTCATGGTCAACCCCGAACCTTCATTGTCA 4895
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 CAAATTTGGACGAGTGGCTGATCTCTTTCTCTTTTTCATGGTCAACCCCGAACCTTCATTGTCA 1440

Qy  4896 ATACTGCAAAAAGAACTGCTGACAATTATGTTTGTTGACTGCGAGCCCAACTACAACTGT 4955
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 ATACTGCAAAAAGAACTGCTGACAATTATGTTTGTTGACTGCGAGCCCAACTACAACTGT 1500

Qy  4956 GTCATCAGTATGGCTATGCTGTCCCAATGACGCACCAGGTGGCAGGGAGCCCGGCTTG 5015
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 GTCATCAGTATGGCTATGCTGTCCCAATGACGCACCAGGTGGCAGGGAGCCCGGCTTG 1560

Qy  5016 GGAAAAAACCTTGTGGGGTTCTGTGGCCTTGAGCGGCGTGACGCCTGTCCTGGCCCCAG 5075
         |||||||||||||||||||||||||||||||| |||  ||||||||||||||||||||
Db  1561 GGAAAAAACCTTGTGTGGGGTTCTGTGGCCGGCGTGACGCCTGTCCTGCGCCCAG 1620

Qy  5076 AGCCCAGCGAGGTGACCCAGATACCAAATGTGCTTCACTGAAGTCAATACTTCTGGACAG 5135
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 AGCCCAGCGAGGTGACCCAGATACCAAATGTGCTTCACTGAAGTCAATACTTCTGGACAG 1680

Qy  5136 CCGCACTCGCTGTGGGTTGAGTGGCTTATCTAGCCATTGACACTTTTGGCG 5195
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 CCGCACTCGCTGTGGGTTGAGTGGCTTATCTAGCCATTGACACTTTTGGCG 1740

Qy  5196 CCACTTGTGTGGGGCGTTGCTGTCTATTACATCAGTCCCTACCGGTGCTACTGTCGCCC 5255
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 CCACTTGTGTGGGGCGTTGCTGTCTATTACATCAGTCCCTACCGGTGCTACTGTCGCCCC 1800

Qy  5256 CAGTGGTTGACGAAGAGAAATCGTGGAGGAGTGCATCATTCATTCCCTTGGCATGCATGG 5315
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 CAGTGGTTGACGAAGAGAAATCGTGGAGGAGTGCATCATTCATTCCCTTGGCATGCATGG 1860

Qy  5316 TGGTTGCTGCAATTGACAAGCTGAAGAGTACAATCACCACAGTCCTTTCACATTGG 5375
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 TGGTTGCTGCAATTGACAAGCTGAAGAGTACAATCACCACAGTCCTTTCACATTGG 1920

Qy  5376 AAACCGCCCTTGAAAAACTTAACACCTTCTTGGCCTCATGCAGCTTACAATCCTTGCTA 5435
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 AAACCGCCCTTGAAAAACTTAACACCTTCTTGGCCTCATGCAGCTTACAATCCTTGCTA 1980

Qy  5436 TCATAGAGTATTGCGTGTGGTTAGTCACTTACCTGACAATCCCTTTGCATCATGCGTGT 5495
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 TCATAGAGTATTGCGTGTGGTTAGTCACTTACCTGACAATCCCTTTGCATCATGCGTGT 2040

Qy  5496 TTGCTTTTCATTGCGGGTATTACTACCCCGCACTACCTCACACACAAAATGTTCCTGTCAT 5555
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 TTGCTTTTCATTGCGGGTATTACTACCCCGCACTACCTCACACACAAAATGTTCCTGTCAT 2100

Qy  5556 TATTTGGAGGCCAATTGCGTCGCGGAACAGCTTCTTGGTACATGGACATCGGTGGGTTTTGTCTTTGACA 5615
         |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 TATTTGGAGGCCAATTGCGTCGCGGAACAGCTTCTTGGTACATGGACATCGGTGGGTTTTGTCTTTGACA 2160

Qy  5616 TGGCCGGGGGCTGCGGGAACAGCTTCTTGGTACATGGACATCGGTGGGTTTTGTCTTTGACA 5675
         |||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 TGGCCGGGGGCTGCGVGGGAACAGCTTCTTGGTACATGGACATCGGTGGGTTTTGTCTTTGACA 2220

Qy  5676 TGCTAGGCGGCTATGCTGCCGGCCTCATCCACTGCTTGCTTGGTTAGTCGACATTTAAATGCTTGATGG 5735
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 TGCTAGGCGGCTATGCTGCCGGCCTCATCCACTGCTTGCTTGGTTAGTCGACATTTAAATGCTTGATGG 2280

Qy  5736 GTGAGTGGCTGCCACTATGGATCAGCTTGCTTGGTTAGTCACTCCGGCGTTCAATCCGGCCG 5795
         ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db  2281 GTGAGTGGCYCACTATGGATCAGCTTGCTTGGTTAGTCACTCCGGCGTTCAATCCGGCCG 2340

Qy  5796 CAGGAGTTGCTGCGGCGTATGCTGTCTTGTCAGTGTTTGCTTTGTCAGAGCAGGGGCCAG 5855
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 CAGGAGTTGTGTGGCGTATGCTGTCTTGTCAGTGTTTGCTTTGTCAGAGCAGGGGCCAG 2400
```

NUMBER OF SEQUENCES: 716
CORRESPONDENCE ADDRESS:
 ADDRESSEE: ABBOTT LABORATORIES   D377/AP6D
 STREET: 100 ABBOTT PARK ROAD
 CITY: ABBOTT PARK
 STATE: IL
 COUNTRY: USA
 ZIP: 60064-3500
COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/488,446
 FILING DATE:
 CLASSIFICATION:
PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/424,550
 FILING DATE:
ATTORNEY/AGENT INFORMATION:
 NAME: POREMBSKI, PRISCILLA E.
 REGISTRATION NUMBER: 33,207
 REFERENCE/DOCKET NUMBER: 5527.PC.01
TELECOMMUNICATION INFORMATION:
 TELEPHONE: 708-937-6365
 TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
 LENGTH: 4268 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
US-08-488-446-80

Query Match   44.2%;   Score 4159;   DB 4;   Length 4268;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 4215; Conservative 20; Mismatches 30; Indels 30; Gaps 2;

Qy 3456 TGGCTCATCCCACAGGCTCTATACACCCAATAACCGTTGACGGCGGCTAATGACCAGGACA 3515
             |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   1 TGGCTCATCCCACAGGCTCCATACACCCAATAACCGTTGACGCGGCTAATGACCAGGACA 60

Qy 3516 TCTATCAACCACCATGTGGAGCTGGGTCCCTTACTCCGGTGCTCTTGCGGGGAGACCAAGG 3575
             ||||||||||||||||||| |||| || ||| |||||||||||||||||||||||||||||
Db   61 TCTATCAACCACCATGTGGAGCTGGGTCCCTTACTCCGGTGCTCTTGCGGGGAGACCAAGG 120

Qy 3576 GGTATCTGGTAACGACTGGGTCATTGGTTGAGGTCAACAAATCCGATGACCCTTATT 3635
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GGTATCTGGTAACGACTGGGTCATTGGTGAGGTCAACAAATCCGATGACCCTTATT 180

Qy 3636 GGTGTGTGTGCCGGGGGCCCTTCCCATGGCTGTTGCCAAGGGTTCTTCAGGTGCCCGATTC 3695
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 GGTGTGTGTGCCGGGGGCCCTTCCCATGGCTGTTGCCAAGGGTTCTTCAGGTGCCCGATTC 240

Qy 3696 TGTGCTCCTCCGGCATGTTATTGGGATGTTGGGATGTTCACCGGGTCGCTAGAAATTCTGGCGGGTTCAG 3755
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 TGTGCTCCTCCGGCATGTTATTGGGATGTTCACCGGGTCGCTAGAAATTCTGGCGGGTTCAG 300

Qy 3756 TCAGTCAGATTAGGGTTAGGCCGTTGAGGCCGTTGGGCTGCTGATACCATCCCAGTACACAGCAC 3815
             ||||||| |||||||| || ||| || ||||||||||||||||||||||||||||||||||||
Db 301 TCGGCCAGATTAGGGTTAGGCCGTTGGGCCGTTGGGCTGCTGATACCATCCCAGTACACAGCAC 360

Qy 3816 ATGCCACTCTTGATACAAAACCTACTGTGCCTAACGAGTATTCAGTGCAAATTTAATTG 3875
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 ATGCCACTCTTGATACAAAACCTACTGTGCCTAACGAGTATTCAGTGCAAATTTAATTG 420

Qy 3876 CCCCCACTGGCAGCGGCAAGTCAACCAAATTACCACTTTTCTTACATGCGAGGAGAGTATG 3935
             ||||||||||||||||||||||||||||||||||||||||||||| : |||||||||||||
Db 421 CCCCCACTGGCAGCGGCAAGTCAACCAAATTACCACTTTTCTTACATGCGAGGRGAAGYATG 480

Qy 3936 AGGTCTTGGTCTCTAAATCCCAGTGTGGCTACAACAGCATCAATGCCAAAGTACATGCACG 3995

Db 3481 GTGACCCGTTGAGAAGGTCCTTCAAMCTCCCTTCTTCCACCACCTGTCTCRCAGTTGGCCA 3540
             ||||||||||||||||||||||||||||| |||||||: |||||||||||||||||||||
Qy 6993 TGCCGATGCCCCTGTTGGGAGCGGCGGTGAGTGTAACCCTTTCACTGCAATTGGATGTGCAA 7052

Db 3541 TGCCGATGCCCCTGTTGGGAGCGGCAGGTGAGTGTAACCCTTTCACTGCAATTGGATGTGCAA 3600
             |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Qy 7053 TGACCGAAACAGGCGGCGGCCCTGATGATTACCCAGTTACCCTCCCAAAAAGAGGAGGTCT 7112

Db 3601 TGACCGAAACARGYGGAGKCCCWSAKRATTACCCAGTTACCCTCCCAAAAAGAGGAGGTCT 3660
             |||||||||| |:|::: |||||||||||||||||||||||||||||||||||||||||
Qy 7113 CTGAATGGTCAGCAGAAAGTTGGTCGACGGCTACAACCGCTTCCAGCTACGTTACTGGCC 7172

Db 3661 CTGAATGGTCAGCAGAACGCGTTGGTCAACGACTACAACCGCTTCCAGCTACGTTACTGGCC 3720
             ||||||||||||||| |||| ||||||||| ||||||||||||||||||||||||||||||
Qy 7173 CCCCGTACCCTAAGATACGGGGAAACGATTCCACTCAGTCAGCCCCCGCAAACGCCTA 7232

Db 3721 CCCCGTACCCTAAGATACGGGGCAAGGATTCCACTCAATCAGCCCCAAACGCCTA 3780
             |||||||||||||||||||||| |||| |||||||||||||||||||||||||||
Qy 7233 CAAAAAGAGAGTTGGGAAAGAGTGAGTTTTCGTGCAGCATGAGCCTACACCTGGACCGACG 7292

Db 3781 CAAAAAGAGAGTTGGGAAAGAGTGAGTTTTCGTGCAGCATGAGCCTACACTTGGACCGACG 3840
             |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Qy 7293 TGATTAGCTTCAAAACTGCTTCTAAAGTTCTGTCTGCAACTCGGGCCATCACTAGTTGGTT 7352

Db 3841 TGATTAGCTTCAAAACTGCTTCTAAAGTTCTGTCTGCAACTCGGGCCATCACTAGTTGGTT 3900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7353 TCCTCAAACAAAGATCATTCGTTGGTTGGTATGTGACTGAGCCGCCGGGGATGCGGAGCTTAGAAAAC 7412

Db 3901 TCCTCAAACAAAGATCATTCGTTGGTTGGTATGTGACTGAGCCGCCGGGGATGCGGAGCTTAGAAAAC 3960
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7413 AAAAAGTCACTATTAATAGACAACCTCGTTCCCCCATCATACCACAAGCAAGTGAGAT 7472

Db 3961 AAAAAGTCACTATTAATAGACAACCTCGTTCCCCCATCATACCACAAGCAAGTGAGAT 4020
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7473 TGGCTAAGGAAAAAGTTCAAAAGTTGTCGGTGTCATGTGGGACTATGTGATGAAGTAGCAG 7532

Db 4021 TGGCTAAGGAAAAAGTTCAAAAGTTGTCGGTGTCATGTGGGACTATGTGATGAAGTAGCAG 4080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7533 CTCACACGCGCCCTCTAAGTCTGCTAAGTCCCACATCACTGGGCCTTCGGGGCACTGATGTTC 7592

Db 4081 CTCACACGCGCCCTCTAAGTCTGCTAAGTCTGCTAAGTCCCACATCACTGGGCCTTCGGGGCACTGA----- 4135

Qy 7593 GTTCTCTGGAGCAGCAGCCCCCGCAAGGCTGTTCTCGGAGCAGTGTGTCGAGGCAGGTGAGA 7652

Db 4136 --------------------TGTTCTCTGGAGCTTGCGCAAGTGTGTCGAGGCAGTGAGA 4173

Qy 7653 TACCGAGTCATTATCGGCAAACTGTGATAGTTCGCAAACTGTGATAGTTCCAAAGGAGGAGGTCTTCGTGAAGACCC 7712

Db 4174 TACCGAGTCATTATCGGCAAACTGTGATAGTTCGCAAACTGTGATAGTTCCAAAGGAGGAGGTCTTCGTGAAGACCC 4233
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 7713 CCCAGAAACCAACAACAAGAAAACCCCCAAGGCTTATC 7747

Db 4234 CCCAGAAACCAACAACAAGAAAACCCCCAAGGCTTATC 4268

RESULT 15
US-08-488-446-80
; Sequence 80, Application US/08488446
; Patent No. 6558898
GENERAL INFORMATION:
 APPLICANT: SIMONS
 APPLICANT: JOHN N.
 APPLICANT: TAMI J. PILOT-MATIAS
 APPLICANT: GEORGE J. DAWSON
 APPLICANT: GEORGE G. SCHLAUDER
 APPLICANT: SURESH M. DESAI
 APPLICANT: THOMAS P. LEARY
 APPLICANT: ANTHONY SCOTT MUERHOFF
 APPLICANT: JAMES C. ERKER
 APPLICANT: SHERI L. BUIJK
 APPLICANT: ISA K. MUSHAHWAR
 TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
 TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE

```
Db                                                              GGAAAAAACCTTGTGGGGTTCTGTCGGCGTTGGACGGCCTGTCCTGCGCCCAG 1620
                                                                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  AGCCCAGCGAGGTGACCAGATCACCAAATGTGCTTCACTGAAGTCAATACTTCTGGGACAG 5135
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  AGCCCAGCGAGGTGACCAGATCACCAAATGTGCTTCACTGAAGTCAATACTTCTGGGACAG 1680

Qy  CCGCACTCGCTGTTGGCGTGGAGTGGCTATCGCTGTATCTAGCCATTGGACACTTTGGCG 5195
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CCGCACTCGCTGTTGGCGTGGAGTGGCTATGGCTATCGCTGTATCTAGCCATTGGACACTTTGGCG 1740

Qy  CACTTGTTGTGCGACGAATGCCATGCCATGCCATCCGTGTCTACTGTCGCCC 5255
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CCACTTGTTGTGCGACGAATGCCATGCCATGCCATCCGTGTCTACTGTCGCCCC 1800

Qy  CAGTGGTTGACGAAGAGAAATCGTGGAGGAGTGCATCATTCCTTGAGGCCA 5315
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CAGTGGTTGACGAAGAGAAATCGTGGAGGAGTGCATCATTCATTCCTTGAGGCCA 1860

Qy  TGGTTGCTGCAATTGACAAGCTGAAGAGTACAATCACCACAACTAGTCCTTCACATTGG 5375
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TGGTTGCTGCAATTGACAAGCTGAAGAGTACAATCACCACAACTAGTCCTTCACATTGG 1920

Qy  AAACCGCCCTTGAAAAACTTAACACCTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTA 5435
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  AAACCGCCCTTGAAAAACTTAACACCTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTA 1980

Qy  TCATAGAGTATTGCTGTGGTTAGTCACTTACCTGACAATCCCTTGCATCATGCGTGT 5495
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TCATAGAGTATTGCTGTGGTTAGTCACTTACCTGACAATCCCTTGCATCATGCGTGT 2040

Qy  TTGCTTTCATTGCGGTATTACTACCCACTACCCACTACCTCACAAGATCAAAATGTTCCTGTCAT 5555
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TTGCTTTCATTGCGGTATTACTACCCACTACCCACTACCTCACAAGATCAAAATGTTCCTGTCAT 2100

Qy  TATTTGGAGGCGCAATTGCGTCCGGGAACAGCTTCGGTACATCGGTGGGTTTGTCTTTGACA 5615
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TATTTGGAGGCGCAATTGCGTCCGGGAACAGCTTCGGTACATCGGTGGGTTTGTCTTTGACA 2160

Qy  TGGCCGGGGCTCGCGGGACAGCTCTTGGTACTTTGACCTACTTTGACCTGGTGGTTTGTCTTTGACA 5675
    ||||||||||||||||:||||||||||||||||||||||||||||||||:||||||||
Db  TGGCCGGGGCTCGYGGGACAGCTCTTGTGTACATGGACATGGACATGGTTGTCTTTGACA 2220

Qy  TGCTAGGCGGCTATGCTGCGCGCCTCATCCACGTGCTTGCTTGTGACATTTAAATGCTTGATGG 5735
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TGCTAGGCGGCTATGCTGCGCGCCTCATCCACTACTGCTTGCTTGTGACATTTAAATGCTTGATGG 2280

Qy  GTGAGTGGCCACTATGGATCAGCTTGCGTGGTTTAGTCTTACTCCGCGTTCAATCCGGCCG 5795
    |||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db  GTGAGTGGCYCACTATGGATCAGCTTGCYCACTATGGATCAGCTTAGTCTTACTCCGCGTTCAATCCGGCCG 2340

Qy  CAGGAGTTGTGGGCGTCTTGTCAGCTTGCGCAATGTTTGCTTTGACAACAGCAGGCCAG 5855
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CAGGAGTTGTGGGCGTCTTGTGTCAGCTTGCGCAATGTTTGCTTTGACAACAGCAGGCCAG 2400

Qy  ATCACTGCCACTGTGACATCCGCAGGAGATACTGGGCATTCTGAGGCATCTACCCC 5915
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  ATCACTGCCACTGTGACATCCGCAGGAGATACTGGGCATTCTGAGGCATCTACCCC 2460

Qy  ACTTTATTGCCACTCGTGACACTGGGCATTCTGAGGCATCTACCTACCCC 5975
    ||||||||||||||||||||||||:||||||||||||||::||||||||
Db  ACTTTATTGCCACTCGTGACATCCGCAGGAGATACTGGGCATTCTGAGGCATCTACCTACCCC 2520

Qy  CCTTGGAGTTCATATCAGCTTGACAATTCAGCTTGCGCAGTATTGGCAATTTCTTTGTGCAATTTGCT 6035
    ||||||||:|||:||||||||||||||||||||:|||||||||||||||||||||||
Db  CCTTGGAGTRTCATATCAGCTTGCGTCTGTCAGCTTGCGCACCCCACGGAGGATTGCG 2580

Qy  GCCTCATTGCTTAGAGATTTGGCAGTATGTGCAATTCTTTGTGATTTGCT 6095
    ||||||||||||:|||:||||||||||||||:||||||||||||||||||
Db  GCCTCATTGCTTRARAGATTTGGCAATTTCTTTGTGATTTGCT 2640

Qy  TTAATGTCCTTAAAGCTGGAGTTCAGACGCATGGTTAACATTCCTTGGTTGTCCTTTCTACA 6155
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TTAATGTCCTTAAAGCTGGAGTTCAGACGCATGGTTAACATTCCTTGGTTGTCCTTTCTACA 2700
```

```
Db                                                              |||||||||||||||||||||||
                                       AGGTCTTGGTCCTAAATCCCAGTGTGCTACAACAGCATCAACAGCCAAAGTACATGCGACG 540

Qy  CGACGTACGGCGTGAATCCAAATTGCTATTTTAATGCAAATGTACCAACACAGGGGCTT 4055
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CGACGTACGGCGTGAATCCAAATTGCTATTTTAATGCAAATGTACCAACACAGGGGCTT 600

Qy  CACTTAGTACAGCACATATGGCATGTACCGGAGCACGTGTTCCCGGACTATGATG 4115
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CACTTAGTACAGCACATATGGCATGTACCGGAGCACGTGTTCCCGGACTATGATG 660

Qy  TAATCATTGTGACGAATGCCATGCCATGCCATCCGTGTTGGGCATTGGAAAGG 4175
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TAATCATTGTGACGAATGCCATGCCATGCCATCCGTGTTGGGCATTGGAAAGG 720

Qy  TCCTAACCGAAGCTCCATCCAAAAATGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCC 4235
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TCCTAACCGAAGCTCCATCCAAAAATGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCC 780

Qy  CTGGAGTAATCCCTACACCACCATGCCAACATAACTGAGATTCAATTAACCTGAAGGCA 4295
    ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db  CTGGAGTAATCCCTACACCACCATGCCAACATAACTGAGATTCAATTAACYGATGAAGGCA 840

Qy  CTATCCCCTTTCATGGAAAAAGATTAAGGAGGAGAAAATCTGAAGAAGGGAGACACCTTA 4355
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CTATCCCCTTTCATGGAAAAAGATTAAGGAGGAGAAAATCTGAAGAAGGGAGACACCTTA 900

Qy  TCTTTGAGGCTACCAAAAAACACTGTGATGAGCTTGCTCAACGAGTTAGCTCGAAAGGGAA 4415
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TCTTTGAGGCTACCAAAAAACACTGTGATGAGCTTGCTCAACGAGTTAGCTCGAAAGGGAA 960

Qy  TAACAGTCTGTCTTCTTACTATAGGGGATGTGACATCTCAAAAATCCCTGAGGGCGACTGTG 4475
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TAACAGTCTGTCTTCTTACTATAGGGGATGTGACATCTCAAAAATGCCTGAGGGCGACTGTG 1020

Qy  TAGTAGTTGCCACTGATGCCTTGTGTACAGGGTACACTGGTACTTTGATTCCGTGTATG 4535
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  TAGTAGTTGCCACTGATGCCTTGTGTACAGGGTACACTGGTACTTTGATTCCGTGTATG 1080

Qy  ACTGCAGCCTCATGGTAGAAGGCACATGCCATGGTCATTTGACCTTACTTTGACCTTTCCACATGG 4595
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  ACTGCAGCCTCATGGTAGAAGGCACATGCCATGGTCATTTGACCTTACTTTCCACATGG 1140

Qy  GTGTTCGTGTGTGCGGGGTTTCAGCAATAGTTAAAGGCCAGCCAGCCCACAGGCC 4655
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  GTGTTCGTGTGTGCGGGGTTTCAGCAATAGTTAAAGGCCAGCCAGCCCCACAGCC 1200

Qy  GTGGGGAGCTGCATATACTACTACTAGAACGGGAGGTTGTACCCCTTCGGGTATGGTTC 4715
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  GTGGGGAGCTGCATATACTACTACTAGAACGGGAGGTTGTACCCCTTCGGGTATGGTTC 1260

Qy  CTGAAATGCAACATTGTTGAAGCCTTGAAGGCATGGTATGGTTTGTCATCAA 4775
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CTGAAATGCAACATTGTTGAAGCCTTGAAGGCATGGTATGGTTTGTCATCAA 1320

Qy  CAGAAGCTCAAAATTTGGACGAGGAGAACTGCTGACAATTGTTTATGTTTGTTGTGACTACAACTGT 4835
    ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db  CAGAAGCTCAAAATTTGGACGAGGAGAACTGCTGACAATTGTTTATGTTTGTTGTGACTACAACTGT 1380

Qy  CAAAATTTGGACGAGGAGAACTGCTGACAATTGTTTATGTTTGTTGTGACTACAACTGT 4895
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  CAAAATTTGGACGAGGAGAACTGCTGACAATTGTTTATGTTTGTTGTGACTACAACTGT 1440

Qy  ATACTGCAAAAGAACTGCTGACAATTGCTCGACAATTGTTGTTGACTGCAGCCCCAACTGT 4955
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  ATACTGCAAAAGAACTGCTGACAATTGCTCGACAATTGTTGTTGACTGCAGCCCCAACTGT 1500

Qy  GTCATCAGTATGGCTATGCTGCTGCTGCTCCCCAATGACGCACCACGGTCGGCAGGGAGCCCGGCTTG 5015
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  GTCATCAGTATGGCTATGCTGCTGCTGCTCCCCAATGACGCACCACGGTCGGCAGGGAGCCCGGCTTG 1560

Qy  GGAAAAAACCTTGTGGGGTTCTGTCGGCGCGTTGGACGGCCTGTCCTGCGCCCAG 5075
    ||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db  GGAAAAAACCTTGTGGGGTTCTGTCGGCGTTGGACGGCCTGTCCTGCGCCCAG 1620
```

```
Qy  6156 GCTGCCAGAAGGGGTACAAGGGCCCCTGGATTGGATCAGGTATGCTTCCAAGCACGCGTGTC 6215
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 GCTGCCAGAAGGGGTACAAGGGCCCCTGGATTGGATCAGGTATGCTTCCAAGCACGCGTGTC 2760

Qy  6216 CATGCGGTGCTGAACTCATCTTTCTGTTGTGAGAATGTGTTTTGCAAAACTTTACAAAGGAC 6275
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 CATGCGGTGCTGAACTCATCTTTCTGTTGTGAGAATGTGTTTTGCAAAACTTTACAAAGGAC 2820

Qy  6276 CCAGAACTTGTTCAAATTACTTGGAGAGGGGCTGTTCCAGTCAACGCTAGGCTGTGTGGGT 6335
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821 CCAGAACTTGTTCAAATTACTTGGAGAGGGGCTGTTCCAGTCAACGCTAGGCTGTGTGGGT 2880

Qy  6336 CGGCTAGACCGGACCCAACTGATTGGACTAGTCTTGTCGTCAATTATGGGCGTTAGGGACT 6395
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881 CGGCTAGACCGGACCCAACTGATTGGACTAGTCTTGTCGTCAATTATGGGCGTTAGGGACT 2940

Qy  6396 ACTGTAAATATGAGAAAATGGGAGATCACATTTTTGTTACAGCAGTATCCTCTCCAAATG 6455
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 ACTGTAAATATGAGAAAATGGGAGATCACATTTTTGTTACAGCAGTATCCTCTCCAAATG 3000

Qy  6456 TCTGTTTCACCCAGGTGCCCCCAACCTTGAGAGCTGCCAGTGCGCCGTTGGACCGGCGTACAGG 6515
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 TCTGTTTCACCCAGGTGCCCCCAACCTTGAGAGCTGCCAGTGCGCCGTTGGACCGGCGTACAGG 3060

Qy  6516 TTCAGTGTTATCTAGGTGAGCCCAAAACTCCTTGGACGACATCTGCTTGCTGTTACGGTC 6575
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 TTCAGTGTTATCTAGGTGAGCCCAAAACTCCTTGGACGACATCTGCTTGCTGTTACGGTC 3120

Qy  6576 CTGACGGTAAGGTAAACTGTAAGCTTCCCTTCCGCGTTGACGGTCACACACCTGGTG 6635
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 CTGACGGTAAGGTAAAACTGTTAAGCTTCCCTTCCGCGTTGACGGACACACACCTGGTG 3180

Qy  6636 TGCGCATGCAACTTAATTTGCGTGATGCACTTGAGACAAATGACTTGTAATTCCACAAACA 6695
         |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db  3181 GTCGCATGCAACTTAATTTGCGTGATGCACTTGAGGCAAATGACTTGTAATTCCATAAACA 3240

Qy  6696 ACACTCCTAGTGATGAAGCCAGTGTCCGCTCCGTTGTTTCAAACAGGAGTTGCGGCGTA 6755
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 ACACTCCTAGTGATGAAGCCAGTGTCCGCTCCGTTCTTGTTTCAAACAGGAGTTGCGGCGTA 3300

Qy  6756 CAAACCAATTGCTTGAGGCAATTTCAGCTGGGTTGACACCACCAAACTGCCAGCCCCCT 6815
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 CAAACCAATTGCTTGAGGCAATTTCAGCTGGGTTGACACCACCAAACTGCCAGCCCCCT 3360

Qy  6816 CC--ATCGAAGAGGTAGTGTAAGAAAGCGCCAGTTCCGGGCAAGAACTTGGTTCGCTTA 6872
         |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 CCCAGATCGAAGAGGTAGTGTAAGAAAGCGCCAGTTCCGGGCAAGAACTTGGTTCGCTTA 3420

Qy  6873 CCTTGCCTCCCCCTCCGAGATCCGTCCCAGAGAGTGTCATGTCCTGAAAGCCTGCAACGAA 6932
         |||||||||||||||||||||||||||::||||||||||||||||||||||||||||||||
Db  3421 CCTTGCCTCCCCCTCCGAGATCCGTCCCAGAGAGTGTCATGTCCTGAAAGCCTGCAACGAA 3480

Qy  6933 GTGACCCGTTAGAAGGTCCTTCAAACCTCCCCTTCTTCACCACCTGTTCTACAGTTGGCCA 6992
         |||||||||||||||||||||:||||||||||||||:|||||||||||||||||||||||
Db  3481 GTGACCCGTTAGAAGGTCCTTCAAMCCTCCCCTTCTTCACCACCTGTTCTRCAGTTGGCCA 3540

Qy  6993 TGCCGATGCCCCGTTGGGAGCGGCGGTGAGTGTAACCCTTTCACTGCGAATTGGATGTGCAA 7052
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541 TGCCGATGCCCCGTTGGGAGCGGCGGTGAGTGTAACCCTTTCACTGCGAATTGGATGTGCAA 3600

Qy  7053 TGACCGAAAACAGGCCCTGATGATTTACCCAGTTACCCGGCTACACCGCTTCCCAAAAAGGAGGTCT 7112
         |||||||||||||||||||||||:::|||||||||||||||||||||||||||||||||||
Db  3601 TGACCGAAACARGYGAGKCCCWSAKRATTTACCCAGTTACCCTTACCCTCCCAAAAAGGAGGTCT 3660

Qy  7113 CTGAATGGTCAGACGAAAGTTGGTCGACGAAAGTTGGTCAACGACTACAACCGCTTCCAGCTACGTTACTGGCC 7172
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661 CTGAATGGTCAGACGAAAGTTGGTCAACGACTACAACCGCTTCCAGCTACGTTACTGGCC 3720

Qy  7173 CCCCGTACCCTAAGATACGGGGAAAGATTCCACTCAGTCAGCCCCCGCCAAACGCCCTA 7232
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721 CCCCGTACCCTAAGATACGGGGCAAGGATTCCACTCAATCAGCCCCAACGCCACCGCCCTA 3780
```

```
Qy  7233 CAAAAAGAAGTTGGGAAAGAGTGAGTTTCGTGCAGCATGAGCTACACCTGACCGACG 7292
         ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  3781 CAAAAAGAAGTTGGGAAAGAGTGAGTTTCGTGCAGCATGAGCTACACTTGACCGACG 3840

Qy  7293 TGATTAGCTTCAAAACTGCTTCTGTCTGCAACTCGGCCATCACTAGTGGTT 7352
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 TGATTAGCTTCAAAACTGCTTCTGTCGTCTGCTGCAACTCGGCCATCACTAGTGGTT 3900

Qy  7353 TCCTCAAACAAAGATCATTGGTGTATGTGACTGGTGCCGGAGCTTAGAAAAC 7412
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 TCCTCAAACAAAGATCATTGGTGTATGTGACTGGTGCCGGAGCTTAGAAAAC 3960

Qy  7413 AAAAAGTCACTATTAATAGACAACCTGTTCCCCCCATCATACCAAGCAAGTGAGAT 7472
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 AAAAAGTCACTATTAATAGACAACCTGTTCCCCCCATCATACCAAGCAAGTGAGAT 4020

Qy  7473 TGGCTAAGGAGAAAAGCTTCAAAAGTTGTCGGTGCATGTGGACTATGAAGTAGCAG 7532
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 TGGCTAAGGAGAAAAGCTTCAAAAGTTGTCGGTGCATGTGGACTATGAAGTAGCAG 4080

Qy  7533 CTCACACGCCCTCTAAGTCTGCTAAGTCCCACATCACTGGCCTTCGGGGCACTGATGTTC 7592
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 CTCACACGCCCTCTAAGTCTGCTAAGTCCCACATCACTGGCCTTCGGGGCACTGA----- 4135

Qy  7593 GTTCTGGAGCAGCCCGGAAGGCTGTTCTGGACTTGCAGAAGTGTGTCGAGGTGAGA 7652
                         --------------TGTTCTGGACTTGCAGAAGTGTGTCGAGGTGAGA 4173
Db  4136

Qy  7653 TACCGAGTTCATTATCGGCAAACTGTGATAGTTCCAAAGGAGGAGGTCTTCGTGAAGACCC 7712
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4174 TACCGAGTTCATTATCGGCAAACTGTGATAGTTCCAAAGGAGGAGGTCTTCGTGAAGACCC 4233

Qy  7713 CCCAGAAACCAACAAAGAAACCCCAAGGCTTATC 7747
         ||||||||||||||||||||||||||||||||||
Db  4234 CCCAGAAACCAACAAAGAAACCCCAAGGCTTATC 4268
```

Search completed: October 30, 2005, 23:28:10
Job time : 991 secs

THIS PAGE BLANK (USPTO)

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:        October 30, 2005, 10:18:21 ; Search time 4375 Seconds
                                     (without alignments)
                                     17743.142 Million cell updates/sec

Title:          US-10-009-002-1
Perfect score:  9399
Sequence:       1 accacaaacaccccagttg.........cccgcctggaattaaaact 9399

Scoring table:  IDENTITY NUC
                Gapop 10.0 , Gapext 1.0

Searched:       9784742 seqs, 412945052 residues

Total number of hits satisfying chosen parameters:      19569484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      Published_Applications NA:*
              1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
              2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
              3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
              4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
              5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
              6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
              7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
              8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
              9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
             10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
             11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
             12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
             13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
             14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
             15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
             16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
             17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
             18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
             19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
             20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
             21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
             22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
             23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
             24: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq2:*
             25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
             26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
             27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
             28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	9399	100.0	9399	9	US-09-742-659-1	Sequence 1, Appli
2	9386.2	99.9	9399	19	US-10-189-359-2	Sequence 2, Appli
3	9059.8	96.4	9143	8	US-08-424-550B-390	Sequence 390, App
4	9059.8	96.4	9143	8	US-08-424-550B-393	Sequence 393, App
5	8692.6	92.5	8912	8	US-08-424-550B-11	Sequence 11, Appl

6	4159	44.2	4268	8	US-08-424-550B-80	Sequence 80, Appl
7	449.8	4.8	479	8	US-08-424-550B-25	Sequence 25, Appl
8	449.8	4.8	479	8	US-08-424-550B-392	Sequence 392, App
c	9	319.6	3.4	337	8	US-08-424-550B-29	Sequence 29, Appl
c	10	319.6	3.4	337	8	US-08-424-550B-159	Sequence 159, App
11	279.4	3.0	281	8	US-08-424-550B-18	Sequence 18, Appl
12	279.4	3.0	281	8	US-08-424-550B-20	Sequence 20, Appl
13	262.4	2.8	5211	17	US-10-328-127-1	Sequence 1, Appli
14	262.4	2.8	5211	17	US-10-328-206-1	Sequence 1, Appli
15	262.4	2.8	5211	26	US-11-140-379-1	Sequence 1, Appli
16	256	2.7	9609	20	US-10-333-449A-33	Sequence 33, Appl
17	254.4	2.7	260	19	US-10-189-359-1	Sequence 1, Appli
18	252.6	2.7	2061	9	US-09-929-955-16	Sequence 16, Appl
19	252.6	2.7	2061	11	US-10-930-591-1	Sequence 1, Appli
20	252.6	2.7	2061	18	US-10-307-047-1	Sequence 16, Appl
21	252.6	2.7	2061	21	US-10-817-591-16	Sequence 16, Appl
22	252.6	2.7	2061	26	US-11-043-808-1	Sequence 1, Appli
23	250.8	2.7	2058	9	US-09-881-654-1	Sequence 1, Appli
24	250.8	2.7	2058	9	US-09-881-239-2	Sequence 2, Appli
25	250.8	2.7	2058	19	US-10-637-323-1	Sequence 1, Appli
26	250.8	2.7	2058	19	US-10-643-853-2	Sequence 2, Appli
27	250.8	2.7	2058	20	US-10-658-782-1	Sequence 1, Appli
28	250.8	2.7	2058	21	US-10-899-715-1	Sequence 1, Appli
29	250.8	2.7	2064	10	US-09-884-456-69	Sequence 69, Appl
30	250.8	2.7	2064	10	US-09-884-455-69	Sequence 69, Appl
31	250.8	2.7	2073	16	US-10-133-133A-5	Sequence 5, Appli
32	250.8	2.7	2523	10	US-09-884-456-85	Sequence 85, Appl
33	250.8	2.7	2523	10	US-09-884-455-85	Sequence 85, Appl
34	250.4	2.7	9379	9	US-09-916-359-1	Sequence 1, Appli
35	250.4	2.7	9401	20	US-10-445-724-1	Sequence 1, Appli
36	247.6	2.6	5965	21	US-10-492-178-10	Sequence 10, Appl
37	244.4	2.6	9646	9	US-09-742-659-3	Sequence 3, Appli
38	244	2.6	1893	16	US-10-214-932-77	Sequence 77, Appl
39	243	2.6	9365	10	US-09-827-688-7	Sequence 7, Appl
40	243	2.6	9416	9	US-09-929-955-13	Sequence 13, Appl
41	243	2.6	9416	14	US-10-966-13	Sequence 13, Appl
42	243	2.6	9416	19	US-10-719-619-13	Sequence 13, Appl
43	243	2.6	9416	21	US-10-817-591-13	Sequence 13, Appl
44	243	2.6	12980	9	US-09-238-076-5	Sequence 5, Appli
45	243	2.6	12980	10	US-09-995-937-5	Sequence 5, Appli

ALIGNMENTS

```
RESULT 1
US-09-742-659-1
; Sequence 1, Application US/09742659
; Patent No. US2001003419A1
; GENERAL INFORMATION:
;   APPLICANT: Hong, Zhi
;   APPLICANT: Butkiewicz, Nancy J.
;   APPLICANT: Zhong, Weidong
;   APPLICANT: Ingravallo, Paul
;   APPLICANT: Wright-Minogue, Jacquelyn
;   APPLICANT: Lau, Johnson Y.
;   APPLICANT: Lemon, Stanley M.
;   TITLE OF INVENTION: Chimeric HCV/GBV-B viruses
;   FILE REFERENCE: ID01116
;   CURRENT APPLICATION NUMBER: US/09/742,659
;   CURRENT FILING DATE: 2000-12-21
;   PRIOR APPLICATION NUMBER: US 60/171,469
;   PRIOR FILING DATE: 1999-12-22
;   NUMBER OF SEQ ID NOS: 16
;   SOFTWARE: Patentin Ver. 2.1
;
; SEQ ID NO 1
;   LENGTH: 9399
;   TYPE: DNA
;   ORGANISM: GB virus-B
US-09-742-659-1

Query Match          100.0%;   Score 9399;   DB 9;   Length 9399;
Best Local Similarity  100.0%;   Pred. No. 0;
```

us-10-009-002-1.rnpb

Matches 9399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1  ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTAGGAGCACCCCCCTAG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  ACCACAAACACTCCCAGTTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG    60

Qy     61  CAGGGCCTGGGGGATTTCCCCGTCTGCGAGAAGGTGAGCCAACCACTTAGTAT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CAGGGCCTGGGGGATTTCCCCGTCTGCGAGAGGGTGAGCCAACCACTTAGTAT        120

Qy    121  GTAGGGCGCGGGACTCATGAGCGTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GTAGGGCGCGGGACTCATGAGCGTCGCGTGATGACAAGCGCCAAGCTTGACTTGGATGGC   180

Qy    181  CCTGATGGGCGTTCATGGGTTCGGTTGGTGGTCGCGGTCCCACGCCCACCA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  CCTGATGGGCGTTCATGGGTTCGGTTGGTGGTCGCGGTCCCACGCCCACCA          240

Qy    241  CCTCCCAGATAGAGCGGGGGACTGTAGGGGAAGACCGGGACACCGGTCACTACCAAGGACG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  CCTCCCAGATAGAGCGGGGGACTGTAGGGGAAGACCGGGACACCGGTCACTACCAAGGACG  300

Qy    301  CAGACCTCTTTTGAGTATCACGCCTCCGGAGTAGTTGGGCAAGCCCACCTATATGTGT
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  CAGACCTCTTTTGAGTATCACGCCTCCGGAGTAGTTGGGCAAGCCCACCTATATGTGT     360

Qy    361  TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGGAT   420

Qy    421  CTGGGAGTTCGTAGACACCGTAGCACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACC
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  CTGGGAGTTCGTAGACACCGTAGCACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACC  480

Qy    481  TGCGCCCAGAACGCGCAAGAACAACAGCAGGCTTCATATCCTGTGTCCATTAAAAC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  TGCGCCCAGAACGCGCAAGAACAACAGCAGGCTTCATATCCTGTGTCCATTAAAAC      540

Qy    541  ATCTGTTGAAAGGGACAACAACGAGCAAAGTCCAGCGCGATGCTCGGCCTCGGCCTGT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  ATCTGTTGAAAGGGACAACAACGAGCAAAGTCCAGCGCGATGCTCGGCCTCGGCCTGT    600

Qy    601  TTACAAAATTGCTGGTATCCAGTCTGCCAGTGCTTGGCTCAGGTGCTTTGCCAGC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  TTACAAAATTGCTGGTATCCAGTCTGCCAGTGCTTGGCTCAGGTGCTTTGCCAGC       660

Qy    661  TCATGGTTGGGGACGCCAAGACCCTCCGCCATAAGTCTCGCAATCTTCGGAATCCTTCTGA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  TCATGGTTGGGGACGCCAAGACCCTCCGCCATAAGTCTCGCAATCTTCGGAATCCTTCTGA  720

Qy    721  TTACCCTTTGGGGTGGATTGGTGATTGCGTGATGTTCACAACTTCACACACCTTAGGGCCTG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  TTACCCTTTGGGGTGGATTGGTGATTGCGTGATGTTCACAACTTCACACACCTTAGGGCCTG 780

Qy    781  GGCAGGAGCGGTCGTTGGTTCGGTTCGGCCGGTGCCGGCTCACTGGTTCGGGGATGGAGTCAA
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  GGCAGGAGCGGTCGTTGGTTCGGTTCGGCCGGTGCCGGCTCACTGGTTCGGGGATGGAGTCAA  840

Qy    841  CTGGGCTACTGGTTCGGTGTCCACCTTTTTGTGGTTATGGTTCCCGCCAATCCTCTGACCAATTG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  CTGGGCTACTGGTTCGGTGTCCACCTTTTTGTGGTTATGGTTCCCGCCAATCCTCTGACCAATTG 900

Qy    901  TCCCTGTAGTGGGGCGCGGTCGTTGGGGTGGGCGGGCCGGGCTCACTGACCAATTG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  TCCCTGTAGTGGGGCGCGGTCGTTGGGGTGGGCGGGCCGGGCTCACTGACCAATTG       960

Qy    961  CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCTTGCCCACTTGCCTACGAGCCTGGTTG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  CTGCCAGCGTAATCAGGTTATCTATTGTTCTCCTTGCCCACTTGCCTACGAGCCTGGTTG   1020

Qy   1021  TGTGATCTGTGCGGACGAGTGCTGTGGGTTCCCGCCAATCCGCCAATCTCACACCCTTCAA
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  TGTGATCTGTGCGGACGAGTGCTGTGGGTTCCCGCCAATCCGCCAATCTCACACCCTTCAA  1080

Qy   1081  TTGGACTGGCACGGACTCCTTCTTGGCTGACCACATTGATTTTGTTATGGGCGCTCTTGT
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  TTGGACTGGCACGGACTCCTTCTTGGCTGACCACATTGATTTTGTTATGGGCGCTCTTGT   1140

Qy   1141  GACCTGTGACGCCCTTGACGTGAGTTGTGTGGTTGTGTGGTTATTAGTCGGTGACTG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  GACCTGTGACGCCCTTGACGTGAGTTGTGTGGTTGTGTGGTTATTAGTCGGTGACTG     1200

Qy   1201  GCTTGTCAGGCACTGGCTTATTCACATAGACTTCAATGGTACTTGTTACCTGA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  GCTTGTCAGGCACTGGCTTATTCACATAGACTTCAATGGTACTTGTTACCTGA         1260

Qy   1261  AGTGCCCACTGAATAGATCCTGGGTTTATCGGGTGGATGGCCGGCAAGGT
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  AGTGCCCACTGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGGATGGCCGGCAAGGT   1320

Qy   1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATTGGCGACTAT
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATTGGCGACTAT        1380

Qy   1381  GTTTAGCAGTGTACACTACCTGGCGGTTGGGCGCTCTGATCTGATCCTATATGCCTCTCGGGGGCAA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381  GTTTAGCAGTGTACACTACCTGGCGGTTGGGCGCTCTGATCTGATCCTATATGCCTCTCGGGGGCAA  1440

Qy   1441  GTGGTATCAGTTGCTCCTAGCGCTTATGCTTTACATAGAAGCGACCTCTGGAAACCCCAT
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441  GTGGTATCAGTTGCTCCTAGCGCTTATGCTTTACATAGAAGCGACCTCTGGAAACCCCAT   1500

Qy   1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTGCGCCTTGATGATACCATGTCC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501  CAGGGTGCCCACTGGATGCTCAATAGCTGAGTTTGCTGCGCCTTGATGATACCATGTCC    1560

Qy   1561  TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561  TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC   1620

Qy   1621  CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621  CAGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG   1680

Qy   1681  TGGCGAGGGGATGGTATGGTTAAATTCAAAAATAACACACATGGGGTTGCTGCGTATTCGCAA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681  TGGCGAGGGGATGGTATGGTTAAATTCAAAAATAACACACATGGGGTTGCTGCGTATTCGCAA  1740

Qy   1741  TGTGCCATCGTACTGCACTATGGCACTGATGCGCAGTGTGGAACGACACTCGCAACACTTA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741  TGTGCCATCGTACTGCACTATGGCACTGATGCGCAGTGTGGAACGACACTCGCAACACTTA  1800

Qy   1801  CGAAGCATGCGGTGTAACACCATGGCTAACAACCGCATGGCACAACGGCTCAGCCCTGAA
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801  CGAAGCATGCGGTGTAACACCATGGCTAACAACCGCATGGCACAACGGCTCAGCCCTGAA   1860

Qy   1861  ATTGGCTATATTACAATACCCTGGGTCTAAAGAAAATGTTTAAACCTCATAATTGGATGTC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861  ATTGGCTATATTACAATACCCTGGGTCTAAAGAAAATGTTTAAACCTCATAATTGGATGTC  1920

Qy   1921  AGGCCATTGTATTTTGAGGGATCAGATCAGTTTTATGACCCTGTGAA
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921  AGGCCATTGTATTTTGAGGGATCAGATCAGTTTTATGACCCTGTGAA             1980

Qy   1981  TTCCACTCTCCTACCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACTCTGTACG
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981  TTCCACTCTCCTACCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACTCTGTACG    2040

Qy   2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGAGACTAGCCACAGG
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041  TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGAGACTAGCCACAGG  2100

Qy   2101  ATTGATCACCAAAGACAAAGCCTGGAAAAAATTATCAGGTCTTATATTCCGCACGGGTGC
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101  ATTGATCACCAAAGACAAAGCCTGGAAAAAATTATCAGGTCTTATATTCCGCACGGGTGC   2160
```

```
Qy  2161  TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTCGGTGCTAATTCTGTTGGGTTGTGTGG  2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TTTGTCTCTTACGGGAGTTACCACCAAGGCCGTCGGTGCTAATTCTGTTGGGTTGTGTGG  2220

Qy  2221  CAGCAAGTATCTTATTTAGGCTCCTCTGTTACTTGTCCCTTGTTTGGGCGCGCTTC      2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  CAGCAAGTATCTTATTTAGGCTCCTCTGTTACTTGTCCCTTGTTTGGGCGCGCTTC      2280

Qy  2281  TGGTTACCCTTTGCGTCTGTCTCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGT      2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  TGGTTACCCTTTGCGTCTGTCTCCATCCCAGTCGTATCTCCAAGCTGGCTGGGATGT      2340

Qy  2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTGCTTGATTTCTTCATCTGTTGCTATCTCCG      2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  TTTGTCTAAAGCTCAAGTAGCTCCTTTGCTTGATTTCTTCATCTGTTGCTATCTCCG      2400

Qy  2401  CTGCAGGCTACGTTATGCTGCCCCTTTAGGGTTTGTGCCCATGGCTGCGGGCGCGTTC      2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  CTGCAGGCTACGTTATGCTGCCCCTTTAGGGTTTGTGCCCATGGCTGCGGGCGCGTTC      2460

Qy  2461  AACTTTCTTTGTTGTGCAGCTGCTGCCCAACCAGATTATGACTGGTGGGTGCGACTGCT   2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  AACTTTCTTTGTTGTGCAGCTGCTGCCCAACCAGATTATGACTGGTGGGTGCGACTGCT   2520

Qy  2521  AGTGGCAGGGTTAGTTTGTGGGGCGCGTAACCGTGGTCACCGCCCATAGCTCTGCTTGT   2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  AGTGGCAGGGTTAGTTTGTGGGGCGCGTAACCGTGGTCACCGCCCATAGCTCTGCTTGT   2580

Qy  2581  AGGTCCTTGGCCTCTCGTAGCGCTTTAACCCTCTTGCATTGGTTACGCCCTGCTTCAGC   2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  AGGTCCTTGGCCTCTCGTAGCGCTTTAACCCTCTTGCATTGGTTACGCCCTGCTTCAGC   2640

Qy  2641  TTTTGATACCGAGATATTGGAGGGCTGACAATACCACCCTGTAGTAGCATTAGTTGTCAT  2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  TTTTGATACCGAGATATTGGAGGGCTGACAATACCACCCTGTAGTAGCATTAGTTGTCAT  2700

Qy  2701  GTCTCGTTTGGCTTCACTTGTTACCTCGCTGCTTAGTTAACTCCTATCT             2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  CGTCTCGTTTGGCTTCACTTGTTACCTCGCTGCTTAGTTAACTCCTATCT            2760

Qy  2761  TTTGGCAACGTTGGGAGAATTGGTTCGTCTCGGAACGTTACACTAAGACCGGAGAGGTTTTTCCT  2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  TTTGGCAACGTTGGGAGAATTGGTTCGTCTCGGAACGTTACACTAAGACCGGAGAGGTTTTTCCT  2820

Qy  2821  TGTGTCTGGTTGTTGTTTCCCCGGTGCGCACATATGACGGCCGTGGTGACTTCTGTGTGTCA   2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  TGTGTCTGGTTGTTGTTTCCCCGGTGCGCACATATGACGGCCGTGGTGACTTCTGTGTGTCA   2880

Qy  2881  CGTAGCTCTTCTATGTTTAACATCCAGTGCAGCATCGTTCTTTGGGACTGACTCTAGGGGT   2940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  CGTAGCTCTTCTATGTTTAACATCCAGTGCAGCATCGTTCTTTGGGACTGACTCTAGGGGT   2940

Qy  2941  TAGGGCCCATAGAATGTTCCTCGGAAGTGTCATGCTTGGTATTCTCATTATGT            3000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941  TAGGGCCCATAGAATGTTCCTCGGAAGTGTCATGCTTGGTATTCTCATTATGT            3000

Qy  3001  TCTTAAGTTTTCCTCTTAGTGTTTGGTGAGAATGGTGTGTTTCTATAAGCACTTGCA       3060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  TCTTAAGTTTTCCTCTTAGTGTTTGGTGAGAATGGTGTGTTTCTATAAGCACTTGCA       3060

Qy  3061  TGGTGATGGTCTTGCCTAATGATTTGCCTGAAACTACCATTGCAAGAGCCATTTCCC       3120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061  TGGTGATGGTCTTGCCTAATGATTTGCCTGAAACTACCATTGCAAGAGCCATTTCCC       3120

Qy  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAAGACGCCTTGCGCGTGTGGGGACAC         3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  TTTTGAAGGCAAGGCAAGGGTCTATAGGAAGACGCCTTGCGCGTGTGGGGACAC         3180

Qy  3181  GGTTGATGGTTTTGCGCCCTTGTTTTCGCAGGGTTGGCTAT                      3240
          |||||||||||||||||||||||||||||||||||||||||
Db  3181  GGTTGATGGTTTTGCGCCCTTGTTTTCGCAGGGTTGGCTAT                      3240

Qy  3241  GCCGCCAGATGGGTGGGCCATTACCGACCTTTTACGCGTCGTCTCTGAACCTGG          3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241  GCCGCCAGATGGGTGGGCCATTACCGACCTTTTACGCGTCGTCTCTGAACGTGG          3300


Qy  3301  CACGCTGTCAGCGATGGCAGTGCAGTGCTAATGACTGGTATAGACCCCGAACTGGACTGGAAC   3360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301  CACGCTGTCAGCGATGGCAGTGCAGTGCTAATGACTGGTATAGACCCCGAACTGGACTGGAAC   3360

Qy  3361  TATCTTCAGATTAGGATCTCTGGCCCACTAGCTACATGGCATACTACATGGGATTTGTTGTGACAACGTGTT  3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361  TATCTTCAGATTAGGATCTCTGGCCCACTAGCTACATGGCATACTACATGGGATTTGTTGTGACAACGTGTT  3420

Qy  3421  GTATACTGCTCACCATGGCAGCAAGGGGCGCCGTTGGCTCATCCCAAGGCTCTATACA      3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421  GTATACTGCTCACCATGGCAGCAAGGGGCGCCGTTGGCTCATCCCAAGGCTCTATACA      3480

Qy  3481  CCCAATAACCGTTGACGCGGCTAATGACCAGCACATCTATCAACCACCATCATCAACCACCATCATCAACCACC  3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481  CCCAATAACCGTTGACGCGGCTAATGACCAGCACATCTATCAACCACCATGTGGAGCTGG  3540

Qy  3541  GTCCCTTACTCGGTGCTCTTGCAGGTGCTCAAGGGGTATCTGTGTGTGCGGGGCCCTTCCCAT  3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541  GTCCCTTACTCGGTGCTCTTGCAGGTGCTCAAGGGGTATCTGTGTGTGCGGGGCCCTTCCCAT  3600

Qy  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCGTGGTCACCGCATAGCTCTGCTTGT         3660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601  ATTGGTTGAGGTCAACAAATCCGATGACCCGTGGTCACCGCATAGCTCTGCTTGT         3660

Qy  3661  GGCTGTTGCCAAGGGGTTCTTCAGGTGCCCGATTCTGTGCTCGTCGTCGTCGTCGTTATTGG  3720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661  GGCTGTTGCCAAGGGGTTCTTCAGGTGCCCGATTCTGTGCTCGTCGTCGTCGTATTGG     3720

Qy  3721  GATGTTCACCGCTGCTAGAAATTCTGGCGGTCAGTCAGTTCAGTAGTTGTCAT          3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721  GATGTTCACCGCTGCTAGAAATTCTGGCGGTCAGTCAGTTCAGTAGTTGTCAT          3780

Qy  3781  GGTGTGTGCTGATACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAACCTAC      3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781  GGTGTGTGCTGATACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAAACCTAC     3840

Qy  3841  TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC    3900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841  TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC    3900

Qy  3901  CAAATTACCACTTCTTACACATGCAGGAGAAGTATGAGGTCTTGGTCTTGGTCCTAAATCCCAGTGT  3960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901  CAAATTACCACTTCTTACACATGCAGGAGAAGTATGAGGTCTTGGTCCTAAATCCCAGTGT  3960

Qy  3961  GGCTACAACAGCATCAATGCCAAAGTCAATGACGCGACGTACGCGGTCATATGGGT       4020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961  GGCTACAACAGCATCAATGCCAAAGTCAATGACGCGACGTACGCGGTGAATCCAAATTG    4020

Qy  4021  CTATTTTAATGCAAATGTACCAACAACAGAGGGGCTTCACTTAGTACAGCACATATGGCAT  4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021  CTATTTTAATGCAAATGTACCAACAACAGAGGGGCTTCACTTACGTACAGCACACATATGGCAT  4080

Qy  4081  GTACCTGACCGGAGCATGTTCCCGGAACTGTTCCCGGAACTACGATGAATGCCATGC      4140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081  GTACCTGACCGGAGCATGTTCCCGGAACTGTTCCCGGAACTACGATGAATGCCATGC      4140

Qy  4141  TACCGATGCAACCACCGTGTTCCCGTGTTGGGCACCGTGTTCCTGGAAAGGTCCATCCAAAAA  4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141  TACCGATGCAACCACCGTGTTCCCGTGTTGGGCACCGTGTTCCTGGAAAGGTCCATCCAAAAA  4200

Qy  4201  TGTTAGGCTAGTTGGTTCTTGCCACGCGTAGGGGTCACCCCCTGGAGTAATCCCTACCACATGC  4260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201  TGTTAGGCTAGTTGGTTCTTGCCACGCGTAGGGGTCACCCCCTGGAGTAATCCCTACACCACATGC  4260

Qy  4261  CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAAGAT   4320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261  CAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCTTTCATGGAAAAAAGAT   4320

Qy  4321  TAAGGAGGAAAAATCTGAAGAAAGGGAGGCACCTTATCTTGAGGCTACCAAAAAACACTG  4380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  4321 TAAGGAGGAAAATCTGAAGAAGGGAGACACCTTATCTTGAGGCTACCAAAAAACACTG 4380
Qy  4381 TGATGAGCTGCTAACCTTAGCTCAGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440
Qy  4441 ATGTGACATCTCAAAAATCCCTGAGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441 ATGTGACATCTCAAAAATCCCTGAGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500
Qy  4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTGCAGCCTCATGGTAGAAGGCAC 4560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTGCAGCCTCATGGTAGAAGGCAC 4560
Qy  4561 ATGCCATGTTGACCTTGACCCTACTTTCACCAGTGGTTCGTGTGTCGGGGGTTTCAGC 4620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561 ATGCCATGTTGACCTTGACCCTACTTTCACCAGTGGTTCGTGTGTCGGGGGTTTCAGC 4620
Qy  4621 AATAGTTAAAGGCCAGCGTAGGGGCCGCCACAGGCCGTGGAGAGCTGGCATATACTACTA 4680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621 AATAGTTAAAGGCCAGCGTAGGGGCCGCCACAGGCCGTGGAGAGCTGGCATATACTACTA 4680
Qy  4681 TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681 TGTAGACGGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740
Qy  4741 CGACGCAGCCAAGGCATGTAGTTGTCATCAACAGAAGCTCAAACTATTCTGGACAC 4800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741 CGACGCAGCCAAGGCATGTAGTTGTCATCAACAGAAGCTCAAACTATTCTGGACAC 4800
Qy  4801 CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT 4860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801 CTATCGCACCCAACCTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCT 4860
Qy  4861 CTTTTCTATGGTCAACCCCGAACCTTCATTTGTCAATACTGCAAAAGAAACTGCTGACAA 4920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861 CTTTTCTATGGTCAACCCCGAACCTTCATTTGTCAATACTGCAAAAGAAACTGCTGACAA 4920
Qy  4921 TTATGTTTGTTGACTGCGAGCCAGCCCAACTACAACTGTGTCATCAGTATGCTATGCTCC 4980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921 TTATGTTTGTTGACTGCGAGCCAGCCCAACTACAACTGTGTCATCAGTATGCTATGCTCC 4980
Qy  4981 CAATGACCGCACCACGGTTGCAGGAGCCCGGCTTGGCGCCTTTGGCGCCGTTGTGTGGGGTTCTGTG 5040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981 CAATGACCGCACCACGGTTGCAGGAGCCCGGCTTGGCGCCTTTGGCGCCGTTGTGTGGGGTTCTGTG 5040
Qy  5041 GCGCTTGACGGCGCGACGGCGTCCTCTGGCCGCCCAGCCCAGCGGAGGTGACCCAGATACCA 5100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041 GCGCTTGACGGCGCGACGGCGTCCTCTGGCCGCCCAGCCCAGCGGAGGTGACCCAGATACCA 5100
Qy  5101 AATGTGCTTCACTGAAGTCAATACTTCATTCCCGGTGCATGGTTGCTGCAATTGACAAGCTGAA 5160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101 AATGTGCTTCACTGAAGTCAATACTTCATTCCCGGTGCATGGTTGCTGCAATTGACAAGCTGAA 5160
Qy  5161 GGCTATGGCTTATCTAGCCATTGACACTTTGACACTTTGCGCGGCCGTTTGCGGGGTTGCTGCTGTG 5220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161 GGCTATGGCTTATCTAGCCATTGACACTTTGACACTTTGCGCGGCCGTTTGCGGGGTTGCTGCTGTG 5220
Qy  5221 TATTACATCAGTCCCTACACGGTGCTACTGGTGTTGACGAAGAGAAATCGT 5280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221 TATTACATCAGTCCCTACACGGTGCTACTGGTGTTGACGAAGAGAAATCGT 5280
Qy  5281 GGAGGAGTGCATGCATCATCAGTGCATGGCATCCCTTGGAGGCCATGGTTGCTGCAATTGACAAGCTGAA 5340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281 GGAGGAGTGCATGCATCATCAGTGCATGGCATCCCTTGGAGGCCATGGTTGCTGCAATTGACAAGCTGAA 5340
Qy  5341 GAGTACAATCACCACCACCAGTCTTTCACATTGGAAACCGCCCCTTGAAAAACTTAACAC 5400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341 GAGTACAATCACCACCACCAGTCTTTCACATTGGAAACCGCCCCTTGAAAAACTTAACAC 5400
Qy  5401 CTTTCTTGGGCCTCATGGCAGCTACAAATCCTTGCTATCATAGAGTATTGCTTGGTTTAGT 5460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401 CTTTCTTGGGCCTCATGGCAGCTACAAATCCTTGCTATCATAGAGTATTGCTTGGTTTAGT 5460
```

```
Qy  5461 CACTTTACCTGACAATCCCTTGCATCATGCGTGTTTGCTTTCATTGCGGGTATTACTAC 5520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461 CACTTTACCTGACAATCCCTTGCATCATGCGTGTTTGCTTTCATTGCGGGTATTACTAC 5520
Qy  5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTGGAGGCGCAATTGCGTCCAA 5580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTGGAGGCGCAATTGCGTCCAA 5580
Qy  5581 GCTTACAGACGCTAGAGGCGCACTGGCGTTCATGGGTCATGATGGCCGGGCTGCGGGAACAGCTCT 5640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581 GCTTACAGACGCTAGAGGCGCACTGGCGTTCATGGGTCATGATGGCCGGGCTGCGGGAACAGCTCT 5640
Qy  5641 TGGTACATGGACATCGGTGGGTTTTGTCTTTGACATGCGTAGGCGGCTAGGCGCTGCCGCCTC 5700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641 TGGTACATGGACATCGGTGGGTTTTGTCTTTGACATGCGTAGGCGGCTAGGCGCTGCCGCCCTC 5700
Qy  5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGGGTGAGTGCCCACTATGGATCAGCT 5760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGGGTGAGTGCCCACTATGGATCAGCT 5760
Qy  5761 TGCTGGTTTAGTCTACTCCGGTCAATCCGGGCCGCGGCCAGGAGTTGTGGGGCGTCTTGTCAGC 5820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761 TGCTGGTTTAGTCTACTCCGGTCAATCCGGGCCGCGGCCAGGAGTTGTGGGGCGTCTTGTCAGC 5820
Qy  5821 TTGTGCAATGTTTGCTTTGACACAGCAGGGCCAGATCACTGGCCCAACAGACTTCTTAC 5880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821 TTGTGCAATGTTTGCTTTGACACAGCAGGGCCAGATCACTGGCCCAACAGACTTCTTAC 5880
Qy  5881 TATGCTTGCTAGGAGCAACACTGCTATGTAATGAGTAATGCTTTATTGCCACTGACATCCG 5940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5881 TATGCTTGCTAGGAGCAACACTGCTATGTAATGAGTAATGCTTTATTGCCACTGACATCCG 5940
Qy  5941 CAGGAAGATACTGGGCATTCGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT 6000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941 CAGGAAGATACTGGGCATTCGGAGGCATCTACCCCCTGGAGTGTCATATCAGCTTGCAT 6000
Qy  6001 CCGTTGGCGTCCACACCCCGACGGAGGATGATTGCGCGGCCTCATTGCTTGCTCTCTAGAGAT 6060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001 CCGTTGGCGTCCACACCCCGACGGAGGATGATTGCGCGGCCTCATTGCTTGCTCTCTAGAGAT 6060
Qy  6061 TTGGCAGTATGTGCAATTTCTTGTGATTGCTTTAATGTCCTTAAAGCTGGAGTTCA 6120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061 TTGGCAGTATGTGCAATTTCTTGTGATTGCTTTAATGTCCTTAAAGCTGGAGTTCA 6120
Qy  6121 GAGCATGGTTAACATTCCTGGTGTCCTTTCTACAGCTGCCAGAAGCTGCCAGAAGGGGCCCC 6180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121 GAGCATGGTTAACATTCCTGGTGTCCTTTCTACAGCTGCCAGAAGCTGCCAGAAGGGGCCCC 6180
Qy  6181 CTGGATTGATCAGGTATGCTCCAAGCACGCTGTCCATGCGGTGCTGACTCATCTTTTC 6240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181 CTGGATTGATCAGGTATGCTCCAAGCACGCTGTCCATGCGGTGCTGACTCATCTTTTC 6240
Qy  6241 TGTTGAGAATGGTTTGCAAAATGTTGTTCAAAATTACTGGAG 6300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241 TGTTGAGAATGGTTTGCAAAATGTTGTTCAAAATTACTGGAG 6300
Qy  6301 AGGGGCTGTCAACGTCAGGTTGGTTGACGAAGACCCAACTGATTG 6360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301 AGGGGCTGTCAACGTCAGGTTGGTTGACGAAGACCCAACTGATTG 6360
Qy  6361 GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGGGAGA 6420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361 GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGGGAGA 6420
Qy  6421 TCACATTTTGTTACGACAGTATCCTCTCCAAATGTCTGTTTTCACCCAGGTGCCCCCAAC 6480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421 TCACATTTTGTTACGACAGTATCCTCTCCAAATGTCTGTTTTCACCCAGGTGCCCCCAAC 6480
Qy  6481 CTTGAGAGCTGCAGCTGCCGTGACGACGGCCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481 CTTGAGAGCTGCAGCTGCCGTGACGACGGCCGTACAGGTTCAGTGTTATCTAGGTGAGCCCAA 6540
```

us-10-009-002-1.rnpb

```
Qy  7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7621 GGACTTGCAGAAGTGTGTCGAGGCAGGTGAGATACCGAGTCATTATCGGCAAACTGTGAT 7680

Qy  7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAAAACAACAAAGAAACCCCCAAG 7740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7681 AGTTCCAAAGGAGGAGGTCTTCGTGAAGACCCCCAGAAAACAACAAAGAAACCCCCAAG 7740

Qy  7741 GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAGATGTACTACGGTCAGGT 7800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7741 GCTTATCTCGTACCCCACCTTGAAATGAGATGTGTTGAGAGATGTACTACGGTCAGGT 7800

Qy  7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTACGGGTTGTAGATCCACGTAC 7860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7801 TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTACGGGTTGTAGATCCACGTAC 7860

Qy  7861 CCGTGTCAAGGCTCTGTTGTGATGGTGGTCACCCGATGCAGTCGGAGCCACATGCGATAC 7920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7861 CCGTGTCAAGGCTCTGTTGTGATGGTGGTCACCCGATGCAGTCGGAGCCACATGCGATAC 7920

Qy  7921 AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTGGAGACAGACATCTACTC 7980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7921 AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTGGAGACAGACATCTACTC 7980

Qy  7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGCATTCACACCATTGGAGGCAGTTATA 8040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7981 AGCAGCTAAACTCAGTGACCAACACCGAGCTGCATTCACACCATTGGAGGCAGTTATA 8040

Qy  8041 CGCTGAGGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTAGGTC 8100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8041 CGCTGAGGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTAGGTC 8100

Qy  8101 TTCCGGCGTCATACTACCTCAAGTTCCAACAGTTTGACCTGCTTGACCTGCTGGTGAAGTAAATGC 8160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8101 TTCCGGCGTCATACTACCTCAAGTTCCAACAGTTTGACCTGCTTGACCTGCTGGTGAAGTAAATGC 8160

Qy  8161 TGCAGCCGAACAGGCTGCATGAAGAACCCTGCTTCCTTATTGCGCCGATGATTGCAC 8220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161 TGCAGCCGAACAGGCTGCATGAAGAACCCTGCTTCCTTATTGCGCCGATGATTGCAC 8220

Qy  8221 CGTAATTGGAAGAGCGCCGGAGCAGAGCGCGACCACAAGCAATGGCGTGTCTTGCTAG 8280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8221 CGTAATTGGAAGAGCGCCGGAGCAGAGCGCGACCACAAGCAATGGCGTGTCTTGCTAG 8280

Qy  8281 CTGGATGAAGGTGATGGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281 CTGGATGAAGGTGATGGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACAGTTTGGA 8340

Qy  8341 AGAATTAACATCATGCTCATCAAATGTTACCTGGAATTACCAAAAGTGGCAAGCCTTA 8400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8341 AGAATTAACATCATGCTCATCAAATGTTACCTGGAATTACCAAAAGTGGCAAGCCTTA 8400

Qy  8401 CTACTTTCTTACAAGATCCTCGTATCCCCCTTGGCAGTGCTCTGCCGAGGGTCTCGGG 8460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8401 CTACTTTCTTACAAGATCCTCGTATCCCCCTTGGCAGTGCTCTGCCGAGGGTCTCGGG 8460

Qy  8461 ATACAACCCCAGTGCTGCGTGGTATCTAATACATCACTACTCCATGTGGGT 8520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8461 ATACAACCCCAGTGCTGCGTGGTATCTAATACATCACTACTCCATGTGGGT 8520

Qy  8521 TAGCCGTGTGTGGCTGTCCATTTCATGGAGCAGATGTCTCTTGAGACAAACTTCCCGA 8580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8521 TAGCCGTGTGTGGCTGTCCATTTCATGGAGCAGATGTCTCTTGAGACAAACTTCCCGA 8580

Qy  8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG 8640

Qy  8641 CATCATTGCTGGTGTGTGCACGGTATTGGGGCTTTCTCCGGTGGTCGCTACACCAACGCTGA 8700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8641 CATCATTGCTGGTGTGTGCACGGTATTGGGGCTTTCTCCGGTGGTCGCTACACCAACGCTGA 8700

Qy  8701 GATCCTCAGAGTTCCCAATCACTAACGAGATGAAGCATGCCCCCCTGCGAGCCCTGCCG 8760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy  6541 AACTCCTTGACGACATCTGCTTGCTGTGTTACGGTCCTGACGGTAAGGGTAAAACTGTTAA 6600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541 AACTCCTTGACGACATCTGCTTGCTGTGTTACGGTCCTGACGGTAAGGGTAAAACTGTTAA 6600

Qy  6601 GCTTCCCTTCCGCGGTTGACGGTCACACACCTGGTGTGCGCATGCATGCAACTTAATTGCGTGA 6660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601 GCTTCCCTTCCGCGGTTGACGGTCACACACCTGGTGTGCGCATGCATGCAACTTAATTGCGTGA 6660

Qy  6661 TGCACTTGAGACAAATGACTGTAATTCCACAAACACACTCTAGTGATGAGCCCAGT 6720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661 TGCACTTGAGACAAATGACTGTAATTCCACAAACACACTCTAGTGATGAGCCCAGT 6720

Qy  6721 GTCCGGCTCTGTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC 6780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721 GTCCGGCTCTGTTTCAAACAGGAGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC 6780

Qy  6781 AGCTGGCGGTTGACACACCACCAAACTGCCAGCCCCTCCATCGAAGAGGTAGTGGTAAGAAA 6840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781 AGCTGGCGGTTGACACACCACCAAACTGCCAGCCCCTCCATCGAAGAGGTAGTGGTAAGAAA 6840

Qy  6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCCTCCGAGATCCGTCCC 6900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCCTCCGAGATCCGTCCC 6900

Qy  6901 AGGAGTGTCATGTCGAAAGCCTGCAACGAAGTGACCCGTTAGAAGGTCCTTCAAACCT 6960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901 AGGAGTGTCATGTCGAAAGCCTGCAACGAAGTGACCCGTTAGAAGGTCCTTCAAACCT 6960

Qy  6961 CCCTCCTTCACCACTGTTCTACAGTTGGCCATGCCAGTCCCTGTTGGGAGCGGGTGA 7020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6961 CCCTCCTTCACCACTGTTCTACAGTTGGCCATGCCAGTCCCTGTTGGGAGCGGGTGA 7020

Qy  7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGGAAACAGGGGCCCTGATGA 7080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021 GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGGAAACAGGGGCCCTGATGA 7080

Qy  7081 TTTACCCAGTTACCCTCCCAAAAAGAGAGGTCTCTGAAATGTCAGACCGAAAGTTGGTCGAC 7140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081 TTTACCCAGTTACCCTCCCAAAAAGAGAGGTCTCTGAAATGTCAGACCGAAAGTTGGTCGAC 7140

Qy  7141 GGCTACAAACCGCTTCCAGCTACGTTACTGGCCCCCCGTACCCTAAGTAACGGGGGAAAGGA 7200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141 GGCTACAAACCGCTTCCAGCTACGTTACTGGCCCCCCGTACCCTAAGTAACGGGGGAAAGGA 7200

Qy  7201 TTCCACTCAGTGCAGTCGGGCGGGGATGCGGGGTCTACAAAAAGAGATTGGGAAAGAGTAGTT 7260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201 TTCCACTCAGTGCAGTCGGGCGGGGATGCGGGGTCTACAAAAAGAGATTGGGAAAGAGTAGTT 7260

Qy  7261 TTCGTGCAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7261 TTCGTGCAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320

Qy  7321 TCTGTCTGCAACTCGGGCCATCAGTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGT 7380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7321 TCTGTCTGCAACTCGGGCCATCAGTAGTGGTTCCTCAAACAAAGATCATTGGTGTATGT 7380

Qy  7381 GACTGAGCGCGCGGGATGCGGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCT 7440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381 GACTGAGCGCGCGGGATGCGGGAGCTTAGAAAACAAAAAGTCACTATTAATAGACAACCTCT 7440

Qy  7441 GTTCCCCCATCATACCACAGCAAGCAGTCGGAGCTTAAGGAAAAAGCTTCAAAAGTTGT 7500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441 GTTCCCCCATCATACCACAGCAAGCAGTCGGAGCTTAAGGAAAAAGCTTCAAAAGTTGT 7500

Qy  7501 CGGTGTCATGTGGGACTATGGGACTCACACGCCCTCTÃAGTCTGCTAAGTC 7560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7501 CGGTGTCATGTGGGACTATGGGACTCACACGCCCTCTCAGTCTGCTAAGTC 7560

Qy  7561 CCACATGCTGGTGTGCAGCGGCCGCCCAGGCGTGTTCT 7620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7561 CCACATGCTGGTGTGCAGCGGCCGCCCAGGCGTGTTCT 7620
```

```
Qy      1 ACCACAAACACTCCAGTTGTTACACTCGCTCGGAATGCTCTCGAGCACCCCCTAG     60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACCACAAACACTCCAGTTGTTACACTCGCTCGGAATGCTCTCGAGCACCCCCTAG     60

Qy     61 CAGGGCGTGGGGGATTCCCCTGCCCGTCTGCGAGAAGGGTGGAGCCAACCACTTAGTAT   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAGGGCGTGGGGGATTCCCCTGCCCGTCTGCGAGAAGGGTGGAGCCAACCACTTAGTAT   120

Qy    121 GTAGGCGGCGGGACTCATGAGCTTGCTCGGCGTGATGACAACGCGCAAGCTTGGATGGC   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTAGGCGGCGGGACTCATGAGCTTGCTCGGCGTGATGACAACGCGCAAGCTTGGATGGC   180

Qy    181 CCTGATGGGCGGTTCATGGGCGGTTCGGTGGTGGTGGTGCAGCCTTAGGCAGCCTCCACGCCACCA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCTGATGGGCGGTTCATGGGCGGTTCGGTGGTGGTGGTGCAGCCTTAGGCAGCCTCCACGCCACCA   240

Qy    241 CCTCCCAGATAGACGCGGCGGCCACTGTAGGGAAGACCGGGGACCGGGGAAGAACCGGGAACG   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTCCCAGATAGACGCGGCGGCCACTGTAGGGAAGACCGGGGACCGGGGAAGAACCGGGAACG   300

Qy    301 CAGACCTCCTTTGAGTATCACGCCTCCGGAGTAGTGGGCAAGCCCACCTATATGTGT   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGACCTCCTTTGAGTATCACGCCTCCGGAGTAGTGGGCAAGCCCACCTATATGTGT   360

Qy    361 TGGGATGGTTGGGGTTAGCCATCCATCACCGTACTGCCTGATAGGGTCCTGGAGGGGAT   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TGGGATGGTTGGGGTTAGCCATCCATCACCGTACTGCCTGATAGGGTCCTGGAGGGGAT   420

Qy    421 CTGGGAGTCTCGTAGACCATGCACATCCGCTGTTATTTCTACTCAAACAAGTCCGTACC   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGGGAGTCTCGTAGACCATGCACATCCGCTGTTATTTCTACTCAAACAAGTCCGTACC   480

Qy    481 TGCGCCCAGAACCGCAAGAACAGCAGCAAAGCGCAAAGTCCAGCGCGAAGTCCAGCGCGAAAGCCTCATTAAAAC   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGCGCCCAGAACCGCAAGAACAGCAGCAAAGCGCAAAGTCCAGCGCGAAGTCCAGCGCGAAAGCCTCATTAAAAC   540

Qy    541 ATCTGTTGAAAGGGACACACGAGCAAAGCGAGCAAAGTCCAGCGCGATGTCGTGGTGACGC   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ATCTGTTGAAAGGGACACACGAGCAAAGCGAGCAAAGTCCAGCGCGATGTCGTGGTGACGC   600

Qy    601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTGCCAGC   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTGCCAGC   660

Qy    661 TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTCGCCAATCTTGGAATCCTTCTGA   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TCATGGTTGGGGACGCCAAGACCCTCGCCATAAGTCTCGCCAATCTTGGAATCCTTCTGA   720

Qy    721 TTACCCTTGTGGGGTGGATTGGTGATGTTACAACTCACACTCACTCTAGTAGGCCGCTGT   780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TTACCCTTGTGGGGTGGATTGGTGATGTTACAACTCACACTCACTCTAGTAGGCCGCTGT   780

Qy    781 GGCAGGAGCGGTCGTTCGACCAGTCTGCCAGTAGTACGCTTGCTGGAGGATGGAGTCAA   840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GGCAGGAGCGGTCGTTCGACCAGTCTGCCAGTAGTACGCTTGCTGGAGGATGGAGTCAA   840

Qy    841 CTGGGCTACTGGTTCGGTTCGGTGTCCACCTTTTTGTGGTGGTATGTCGTCTATCTTTGGCCTG   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTGGGCTACTGGTTCGGTTCGGTGTCCACCTTTTTGTGGTGGTATGTCGTCTATCTTTGGCCTG   900

Qy    901 TCCCTGTAGTGGGGCGCGGGGGTCACTGACCCTGACCCCCAGACACAAATACCACAATTG   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCCCTGTAGTGGGGCGCGGGGGTCACTGACCCTGACCCCCAGACACAAATACCACAATTG   960

Qy    961 CTGCCAGCGGAATCAGCGTCGGGTTATCTATTGTTCTCCTTCCACTTGCCTACACGAGCCTGGTTG  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CTGCCAGCGGAATCAGCGTCGGGTTATCTATTGTTCTCCTTCCACTTGCCTACACGAGCCTGGTTG  1020

Qy   1021 TGTGATCGTGCGCGACGACTCCTTCTTGGCTGCTGGTTCCCGCCAATCCTCCACCCTTCCAA  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TGTGATCGTGCGCGACGACTCCTTCTTGGCTGCTGGTTCCCGCCAATCCTCCACCCTTCCAA  1080

Qy   1081 TTGGACTGCACCGGACTCCTTCTTGGCTGACCACATTGATTTTGTTATGGGCGCTCTTGT  1140
```

```
Db   8701 GATCCTCAGAGTTCCAATCACTAACAGACAGACATGACCATGCCCCCTGCGAGCCTGCG  8760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   8761 AAAGAAAGCCAGGGCGGTCCTGCCAGCGCCAAGAGCGTGGCGGACACGCGCCAAAATT  8820
Db   8761 AAAGAAAGCCAGGGCGGTCCTGCCAGCGCCAAGAGCGTGGCGGACACGCGCCAAAATT  8820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   8821 GGCTCGCTCCTTCCTTCCTTCTCGGCATCGCTACATCTAGACCTCTACCAGATTTGGATAAGACGAG  8880
Db   8821 GGCTCGCTCCTTCCTTCCTTCTCGGCATCGCTACATCTAGACCTCTACCAGATTTGGATAAGACGAG  8880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   8881 CGTGGCTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGAGGGGGATGTGTT  8940
Db   8881 CGTGGCTCGGTACACCACTTTCAATTATTGTGATGTTTACTCCCCGAGGGGGATGTGTT  8940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   8941 TATTACACCACAGAGAGAGATTGCAGAAGTTCCTGTGAAGTATTCCTGTGTCATTGTTT  9000
Db   8941 TATTACACCACAGAGAGAGATTGCAGAAGTTCCTGTGAAGTATTCCTGTGTCATTGTTT  9000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   9001 TGCCCTAGGGGCTCATTGCGTGTTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA  9060
Db   9001 TGCCCTAGGGGCTCATTGCGTGTTGGATTAGCCATCAGCTGAACCCCCAAATTCAAAATTAA  9060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   9061 CTAACAGTTTTTTTTTTTTTTTTAGGGCAGCGGAACAGAGGGAGACCCCCGGGC  9120
Db   9061 CTAACAGTTTTTTTTTTTTTTTTAGGGCAGCGGAACAGAGGGAGACCCCCGGGC  9120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   9121 TTAACGACCCCGCGATGTGAGTTTGGCGACCATGGTTGGATCAGAACCGTTTCGGGTGAA  9180
Db   9121 TTAACGACCCCGCGATGTGAGTTTGGCGACCATGGTTGGATCAGAACCGTTTCGGGTGAA  9180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   9181 GCCATGTGTCGAAGGGATGAGGGGATGACCTCCCTTCTGGCTCCATCCACAAAAACCGTCTCGGGTGG  9240
Db   9181 GCCATGTGTCGAAGGGATGAGGGGATGACCTCCCTTCTGGCTCCATCCACAAAAACCGTCTCGGGTGG  9240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   9241 GTGAGGAGTCCTGGCTGTGTGTGGGAAGCAGTCAGTAGTACGCTTGCTGCAGAGCGTAGTGCACCCCCGGGTTTTG  9300
Db   9241 GTGAGGAGTCCTGGCTGTGTGTGGGAAGCAGTCAGTAGTACGCTTGCTGCAGAGCGTAGTGCACCCCCGGGTTTTG  9300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   9301 CTCACGACGTATTTGTCCGCTGCGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTG  9360
Db   9301 CTCACGACGTATTTGTCCGCTGCGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTG  9360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   9361 TTCCAAGCGGAGGGCAACCCCCCGCCTTGGAATTAAAAACT  9399
Db   9361 TTCCAAGCGGAGGGCAACCCCCCGCCTTGGAATTAAAAACT  9399
```

```
Db   1081 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TTGGACTGGCACGGACTCCTTCTTGGCTGACCACATTGATTTGTTATGGGCGCTCTTGT 1140
Qy   1141 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GACCTGTGACGCCCTTGACATTGGTGAGTTGTGTGGTGCGTGTATTAGTCGGTGACTG 1200
Db   1141 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GACCTGTGACGCCCTTGACATTGGTGAGTTGTGTGGTGCGTGTATTAGTCGGTGACTG 1200

Qy   1201 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GCTTGTCAGGGCACTGGCTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGA 1260
Db   1201 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GCTTGTCAGGGCACTGGCTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGA 1260

Qy   1261 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          AGTGCCCACTGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGATGGCCGGCAAGGT 1320
Db   1261 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          AGTGCCCACTGAATAGATCCTGGGTTCCTAGGGTTTATCGGGTGATGGCCGGCAAGGT 1320

Qy   1321 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAGTACCATACGCTATTGCGACTAT 1380
Db   1321 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAGTACCATACGCTATTGCGACTAT 1380

Qy   1381 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GTTTAGCAGTGTACACTACCTGGCGGTTGGCGCTCTGATCTCATGCCTCTCGGGGCAA 1440
Db   1381 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GTTTAGCAGTGTACACTACCTGGCGGTTGGCGCTCTGATCTCATGCCTCTCGGGGCAA 1440

Qy   1441 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GTGGTATCAGTTGCTCCTAGCGCTTATGCTTTACATAGAAGCGACCTCTGAAAACCCAT 1500
Db   1441 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          GTGGTATCAGTTGCTCCTAGCGCTTATGCTTTACATAGAAGCGACCTCTGAAAACCCTAT 1500

Qy   1501 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CAGGGGTGCCCACTGAATGCTCAATAGCTGAGTTTGCTCGCCTTTGATGATACCATGTCC 1560
Db   1501 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CAGGGGTGCCCACTGAATGCTCAATAGCTGAGTTTGCTCGCCTTTGATGATACCATGTCC 1560

Qy   1561 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TTGCCACTCTTATTTGAGTGAGAATGCTCAGAAGTCATTGTTGTACAGTCCAAGTGGAC 1620
Db   1561 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TTGCCACTCTTATTTGAGTGAGAATGCTCAGAAGTCATTGTTGTACAGTCCAAGTGGAC 1620

Qy   1621 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CAGGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTG 1680
Db   1621 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CAGGGCCTATCACTCTAGAGTATAACAACTCCATATCTTGGTACCCCTATACAATCCCTGG 1680

Qy   1681 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TGCGAGGGGATGTATGGCTAAATTCAAAATAACACCATGGGGTTGCTGCCGTATTCGCAA 1740
Db   1681 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TGCGAGGGGATGTATGGCTAAATTCAAAATAACACCATGGGGTTGCTGCCGTATTCGCAA 1740

Qy   1741 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TGTGCCATCGTACTGACTACTGGCACTGATGCAGTGTGGAACGAACACTCGCAACACTTA 1800
Db   1741 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TGTGCCATCGTACTGACTACTGGCACTGATGCAGTGTGGAACGAACACTCGCAACACTTA 1800

Qy   1801 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CGAAGCATGCGGTGTAACACCGATGCTAACACCGGCTGTCAGCCCTGAA 1860
Db   1801 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CGAAGCATGCGGTGTAACACCGATGCTAACACCGGCTGTCAGCCCTGAA 1860

Qy   1861 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          ATTGGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC 1920
Db   1861 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          ATTGGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC 1920

Qy   1921 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          AGGCCATTTGTATTTGAGGGATCAGATATAGTTTACTTTTATGACCCTGTGAA 1980
Db   1921 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          AGGCCATTTGTATTTGAGGGATCAGATATAGTTTACTTTTATGACCCTGTGAA 1980

Qy   1981 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TTCCACTCTCTCCACCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCGGTACG 2040
Db   1981 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TTCCACTCTCTCCACCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCGGTACG 2040

Qy   2041 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC 2100
Db   2041 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGGTCTTATATTCCGCCACGGGTGC 2100

Qy   2101 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TTTGTCTCTTACGGAGTTACCACCAGGCCGTGGTGCTAATTCTGTTGTGGGGTTGTGTGG 2160
Db   2101 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          TTTGTCTCTTACGGAGTTACCACCAGGCCGTGGTGCTAATTCTGTTGTGGGGTTGTGTGG 2160
```

```
Db   2161 TTTGTCTCTTACGGAGTTACCACCAGGCCGTGGTGCTAATTCTGTTGTGGGGTTGTGTGG 2220
Qy   2221 CAGCAAGTACGTATCCTTATTTTAGCCTACCCTGTTGTGTCCCCTTGTTTTGGGGCGCGCCTTC 2280
Db   2221 CAGCAAGTACGTATCCTTATTTTAGCCTACCCTCGTGTCCCATCCCGTGTTTGGGGCGCGCCTTC 2280

Qy   2281 TGGTTACCCCTTTGCGTCCTGTGCTCCATCCCAGTCGTATCTCCAAGCTGGCTGGGGATGT 2340
Db   2281 TGGTTACCCCTTTGCGTCCTGTGCTCCATCCCAGTCGTATCTCCAAGCTGGCTGGGGATGT 2340

Qy   2341 TTTGTCTAAAGCTCAAGTAGCTCCCTTTGCTTGATTTCTTCATCTGTGTCATCTCCG 2400
Db   2341 TTTGTCTAAAGCTCAAGTAGCTCCCTTTGCTTGATTTCTTCATCTGTGTCATCTCCG 2400

Qy   2401 CTGCAGGCTACGTTATGCTCCCATGGCTTGTGCCCATGGCTTGCCCCCT 2460
Db   2401 CTGCAGGCTACGTTATGCTCCCATGGCTTGTGCCCATGGCTTGCCCCCT 2460

Qy   2461 AACTTTCTTCTTGTCGCTGCAGCTGCCCCAACCAGATTATGACTGGTGGTGCGACTGCT 2520
Db   2461 AACTTTCTTCTTGTCGCTGCAGCTGCCCCAACCAGATTATGACTGGTGGTGCGACTGCT 2520

Qy   2521 AGTGGCAGGGTTAGTTTGTGTGGGGCCGGCCGTAACCGTGTCACCGCATAGCTCTGCTTGT 2580
Db   2521 AGTGGCAGGGTTAGTTTGTGTGGGGCCGGCCGTAACCGTGTCACCGCATAGCTCTGCTTGT 2580

Qy   2581 AGGTCCTTGGCCTCTGAGGCCTTTAACCCTCTTGTCATTTGGTTGTTACGCCTTCAGC 2640
Db   2581 TTTGTCTAAAGCTCAAGTAGCTCCCTTTGCTTGATTTCTTCATCTGTGTCATCTCCG 2640

Qy   2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGCATTAGTTGTCAT 2700
Db   2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGCATTAGTTGTCAT 2700

Qy   2701 GTCTCGTTTTGGCTTCTTGCTCACTTGTTAGTTGCTTAGTTGCTTAGTTAACTCCTATCT 2760
Db   2701 GTCTCGTTTTGGCTTCTTGCTCACTTGTTAGTTGCTTAGTTGCTTAGTTAACTCCTATCT 2760

Qy   2761 TTGGCAACGTTGGGAGAATTGGGAACGTTGGTTTGGGCAACGTTGGTTTTCCT 2820
Db   2761 TTGGCAACGTTGGGAGAATTGGGAACGTTGGTTTGGGCAACGTTGGTTTTCCT 2820

Qy   2821 TGTGCTGGTTTGTTTCCCCGGTGGACATATGACGCGGCTGGTGACTTTCTGTGTGTCA 2880
Db   2821 TGTGCTGGTTTGTTTCCCCGGTGGACATATGACGCGGCTGGTGACTTTCTGTGTGTCA 2880

Qy   2881 CGTAGCTCTTCTATGTTTAACATCCAGTGCAGCATCGTCTTGGGACTGACTCTAGGGT 2940
Db   2881 CGTAGCTCTTCTATGTTTAACATCCAGTGCAGCATCGTCTTGGGACTGACTCTAGGGT 2940

Qy   2941 TAGGGCCCATAGAAATGTTGGTGCGTCTCGGAAAGTGTCATGCTTGGTATTCTCATTATGT 3000
Db   2941 TAGGGCCCATAGAAATGTTGGTGCGTCTCGGAAAGTGTCATGCTTGGTATTCTCATTATGT 3000

Qy   3001 TCTTAAGTTTTTCCCTCTTCCCTCTTAGTGTTTGGTGAGAATGTGTGTTTTCTCATAAGCACTTGCA 3060
Db   3001 TCTTAAGTTTTTCCCTCTTCCCTCTTAGTGTTTGGTGAGAATGTGTGTTTTCTCATAAGCACTTGCA 3060

Qy   3061 TGGTGATGTCTTGCCTAATGATTTTGCCTCGGGCGACTACCATTTTTCCC 3120
Db   3061 TGGTGATGTCTTGCCTAATGATTTTGCCTCGGGCGACTACCATTTTTCCC 3120

Qy   3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACGCTTGGCGTGTGGGGACAC 3180
Db   3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGACGCTTGGCGTGTGGGGACAC 3180

Qy   3181 GGTTGATGGTTTGCCCGGTTGTTTTCGGCAGGGTTGGCTAT 3240
Db   3181 GGTTGATGGTTTGCCCGGTTGTTTTCGGCAGGGTTGGCTAT 3240

Qy   3241 GCCGCCAGATGGGTGGGCCATTACCGACCTTTTACGCTGCAGTGTCTCTCGAACGTGG 3300
Db   3241 GCCGCCAGATGGGTGGGCCATTACCGACCTTTTACGCTGCAGTGTCTCTCGAACGTGG 3300
```

```
Qy  3301 CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC 3360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTGGAAC 3360

Qy  3361 TATCTTCAGATTAGGGATCTCTGCCACTACTGTCTGTGCCACTGGGGATTTGTTGTTGTGACAACGTGTT 3420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 TATCTTCAGATTAGGGATCTCTGCCACTACTGTCTGTGCCACTGGGGATTTGTTGTTGTGACAACGTGTT 3420

Qy  3421 GTATACTGCTCACCATGGCAGCAGGGGCGCCCGTTGCTCATCCCACAGGCTCTATACA 3480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421 GTATACTGCTCACCATGGCAGCAGGGGCGCCCGTTGCTCATCCCACAGGCTCTATACA 3480

Qy  3481 CCCAATAACCGTTGACGCGTGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG 3540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481 CCCAATAACCGTTGACGCGTGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG 3540

Qy  3541 GTCCCTTACTCGGTGCTCTTGCGGGGAGCAAGGGGTATCTGGTAACAGACGACTGGGGTC 3600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541 GTCCCTTACTCGGTGCTCTTGCGGGGAGCAAGGGGTATCTGGTAACAGACGACTGGGGTC 3600

Qy  3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGCGGGGCCCTTCCCAT 3660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGCGGGGCCCTTCCCAT 3660

Qy  3661 GGCTGGTGTGCCAAGGGGTTCTTCAGGTGCTCCTCCGGGCATGTTATTGG 3720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661 GGCTGGTGTGCCAAGGGGTTCTTCAGGTGCTCCTCCGGGCATGTTATTGG 3720

Qy  3721 GATGTTCACCGCTGCTAGAAATTCTGGCGGTTCAGTCAGATTAGGGTTAGGCCGTT 3780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721 GATGTTCACCGCTGCTAGAAATTCTGGCGGTTCAGTCAGATTAGGGTTAGGCCGTT 3780

Qy  3781 GGTGTGTGCTGGATACCATCCCAGTACACAGACATGCCACTCTTGATACAAACCTAC 3840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781 GGTGTGTGCTGGATACCATCCCAGTACACAGACATGCCACTCTTGATACAAACCTAC 3840

Qy  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTAGGCTGCAAGTCAAC 3900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTAGGCTGCAAGTCAAC 3900

Qy  3901 CAAATTACCACTTCTTACATGGCAAGGAAGAGTATGAGGTCTTGGTCCTAAATCCCAGTGT 3960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 CAAATTACCACTTCTTACATGGCAAGGAAGAGTATGAGGTCTTGGTCCTAAATCCCAGTGT 3960

Qy  3961 GGCTACAACAGCATCAATGCCAAAGTACGCGACGTACGCGATGTGAATCAAATTG 4020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 GGCTACAACAGCATCAATGCCAAAGTACGCGACGTACGCGATGTGAATCAAATTG 4020

Qy  4021 CTATTTTAATGCAAATGTACCACACAGGGGCTTCACTTACGTAACAGCACATATGGCAT 4080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 CTATTTTAATGCAAATGTACCACACAGGGGCTTCACTTACGTAACAGCACATATGGCAT 4080

Qy  4081 GTACCTGACCGGAGCATGTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCCATGC 4140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 GTACCTGACCGGAGCATGTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCCATGC 4140

Qy  4141 TACCGGATGCAACCACCGTGTTGGGCATTGGGAAGGTCCTAACCGAAGCTCCATCCAAAAA 4200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 TACCGGATGCAACCACCGTGTTGGGCATTGGGAAGGTCCTAACCGAAGCTCCATCCAAAAA 4200

Qy  4201 TGTTAGGCTAGTGGTTGTTGCCACGGCTACCCCCCTGGAGTAATCCCTGGAATCACCACATGC 4260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201 TGTTAGGCTAGTGGTTGTTGCCACGGCTACCCCCCTGGAGTAATCCCTGGAATCACCACATGC 4260

Qy  4261 CAACATAAACTGAGATTCAATTGAAGGCACTTATCTTCATGGAAAAAAGAT 4320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261 CAACATAAACTGAGATTCAATTGAAGGCACTTATCTTCATGGAAAAAAGAT 4320

Qy  4321 TAAGGAGAAAATCTGAAGAAAGGGAGACACCTTATCTTTGAGGCTACCAAAAAAACACTG 4380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321 TAAGGAGAAAATCTGAAGAAAGGGAGACACCTTATCTTTGAGGCTACCAAAAAAACACTG 4380
```

```
Qy  4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381 TGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTACTATAGGGG 4440

Qy  4441 ATGTGACATCTCAAAAATCCCTGAGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441 ATGTGACATCTCAAAAATCCCTGAGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500

Qy  4501 TACAGGGTACACGTGGACTTTGATTCCGTGTATGACTGACTGCAGCGTCATGGTAGAGGCAC 4560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501 TACAGGGTACACGTGGACTTTGATTCCGTGTATGACTGACTGCAGCGTCATGGTAGAGGCAC 4560

Qy  4561 ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGGTGTCGTGTGTCGTGTGTGGGTTTCAGC 4620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561 ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGGTGTCGTGTGTCGTGTGTGGGTTTCAGC 4620

Qy  4621 AATAGTTAAAGGCCAGCGTAGGGGCTGGGGCCGCACGCCGTGGGAGAGCTGGCATATACTACTA 4680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621 AATAGTTAAAGGCCAGCGTAGGGGCTGGGGCCGCACGCCGTGGGAGAGCTGGCATATACTACTA 4680

Qy  4681 TGTAGACGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681 TGTAGACGGAGTTGTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGCCTT 4740

Qy  4741 CGACGCAGCCAAGGCATGTAGTGTTGTTGTCATCAACAGAGCTCAAACTATTCTGGACAC 4800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741 CGACGCAGCCAAGGCATGTAGTGTTGTTGTCATCAACAGAGCTCAAACTATTCTGGACAC 4800

Qy  4801 CTATCGCACCAACCTGGGTTACCTGGTTACCTGCGATAGGAGCAAATTTGACAGTGGGCTGATCT 4860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801 CTATCGCACCAACCTGGGTTACCTGGTTACCTGCGATAGGAGCAAATTTGACAGTGGGCTGATCT 4860

Qy  4861 CTTTTCTATGGTCAACCCCGAACCTTCATTGTCAATACTGCTGACAA 4920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861 CTTTTCTATGGTCAACCCCGAACCTTCATTGTCAATACTGCTGACAA 4920

Qy  4921 TTATGTTTTGTTGACTGCAGCCACCTACAACTGTGTCATCAGTCAGTGCTCC 4980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921 TTATGTTTTGTTGACTGCAGCCACCTACAACTGTGTCATCAGTCAGTGCTCC 4980

Qy  4981 CAATGACGGCACCACCGGTGGCACACCGGTTGGGAAAAAACCTTGTGTGGGGTTCTGTG 5040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981 CAATGACGGCACCACCGGTGGCACACCGGTTGGGAAAAAACCTTGTGTGGGGTTCTGTG 5040

Qy  5041 GCGCTTGACGCGGCGCTGACGCGTCCTGGCCCCAGAGCCCAGCGAGGTGACCAGATACCA 5100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041 GCGCTTGACGCGGCGCTGACGCGTCCTGGCCCCAGAGCCCAGCGAGGTGACCAGATACCA 5100

Qy  5101 AATGTGCTTCACTGAAGTCAATACTTCTGGGGACAGCCCGTGTTGGGCGTTGGAGT 5160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101 AATGTGCTTCACTGAAGTCAATACTTCTGGGGACAGCCCGTGTTGGGCGTTGGAGT 5160

Qy  5161 GGCTATGGCTTATCTAGCCATTGACACTTTGCGCGCCCACTTGTGTGTGGCGCGTTGCTGCGGTC 5220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161 GGCTATGGCTTATCTAGCCATTGACACTTTGCGCGCCCACTTGTGTGTGGCGCGTTGCTGCGGTC 5220

Qy  5221 TATTACATCAGTCCCTACCGGTCCTACTGGTCGCCCGGTGGTTGACGAAGAAATCGT 5280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221 TATTGCATCAGTCCCTACCGGTGCTACTGGTGCTACTGGTTGACGAAGAAATCGT 5280

Qy  5281 GGAGGAGTGTGCATCATTCATTCCCTGGAGGCCATGTGGGCCATGTGGCAATTGACAAGCTGAA 5340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281 GGAGGAGTGTGCATCATTCATTCCCTGGAGGCCATGTGGGCCATGTGGCAATTGCTGCAATCGATAAGCTGAA 5340

Qy  5341 GAGTACAATCACCACCACTAGTCCTTACATTGGAAACCGCCCTTGAAAAACTTAACAC 5400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341 GAGTACAATCACCACCACTAGTCCTTACATTGGAAACCGCCCTTGAAAAACTTAACAC 5400

Qy  5401 CTTTCTTGGGCCTTCATGCAATCCCTTTGCATCATGAGCTACTCATCATAGAGTATTGCTGTGTTTAGT 5460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401 CTTTCTTGGGCCTTCATGCAATCCCTTTGCATCATGAGCTACTCATCATAGAGTATTGCTGTGTTTAGT 5460

Qy  5461 CACTTTACCTGACAATCCCTTTGCATCATGCGTGTTTGCTTTCATTGCGGGTATTACTAC 5520
```

Db   5461 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 5520
        CACTCTTTACCTGACAAATCCCTTTGCATCATGCGTGTTTGCTTTCATTGCGGGTTATTACTAC

Qy   5521 CCCACTACCTTCACAGAGATCAAAAATGTTCCTGTCATTATTTGGAGGCGCAATTGCGTCCAA 5580
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5521 CCCACTACCTTCACAGAGATCAAAAATGTTCCTGTCATTATTTGGAGGCGCAATTGCGTCCAA 5580

Qy   5581 GCTTACAGACGCTAGAGGCGACTGGCGTTCATGATGGCCCGGGGCTGCGGGAACGAGCAGCTCT 5640
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5581 GCTTACAGACGCTAGAGGCGACTGGCGTTCATGATGGCCCGGGGCTGCGGGAACGAGCAGCTCT 5640

Qy   5641 TGGTACATGGACATCGGTGGGTTTGTCTTGACATGCTAGGCGGCTATGCTGCTGCCGCCTC 5700
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5641 TGGTACATGGACATCGGTGGGTTTGTCTTGACATGCTAGGCGGCTATGCTGCTGCCGCCTC 5700

Qy   5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGGCCCACTATGGATCAGCT 5760
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5701 ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGGCCCACTATGGATCAGCT 5760

Qy   5761 TGCTGGTTTAGTCTACTCCGCGTTCAATCCGGCCGGCCGGGCGTTGTGGCGTCTTGTCAGC 5820
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5761 TGCTGGTTTAGTCTACTCCGCGTTCAATCCGGCCGGCCGGGCGTTGTGGCGTCTTGTCAGC 5820

Qy   5821 TTGTGCCAATGTTGCTTTGACAACAGCAGGCCAGATCACTCGGCCCAACAGACTTCTTAC 5880
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5821 TTGTGCCAATGTTGCTTTGACAACAGCAGGCCAGATCACTCGGCCCAACAGACTTCTTAC 5880

Qy   5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGACATCCG 5940

Qy   5941 CAGGAAGATACTGGGCATTCTGGAGCATTCGGAGTGTCATATCACCCCGAGTGTCTTGCAT 6000
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5941 CAGGAAGATACTGGGCATTCTGGAGCATTCGGAGTGTCATATCACCCCGAGTGTCTTGCAT 6000

Qy   6001 CCGTTGGCTGGCACACCCGACGAGGATGATTGCGCGCTCATTGGGTCCTCAGTGCTAGAGAT 6060
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6001 CCGTTGGCTGGCACACCCGACGAGGATGATTGCGCGCTCATTGGGTCCTCAGTGCTAGAGAT 6060

Qy   6061 TTGGCAGTATGTGTGCAATTTCTTTGTGATTTGCTTAAATGTCCTTAAAGCTGGAGTTCA 6120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6061 TTGGCAGTATGTGTGCAATTTCTTTGTGATTTGCTTAAATGTCCTTAAAGCTGGAGTTCA 6120

Qy   6121 GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAGGGGTACAAGGGCCC 6180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6121 GAGCATGGTTAACATTCCTGGTTGTCCTTTCTACAGCTGCCAGAGGGGTACAAGGGCCC 6180

Qy   6181 CTGGGATTGGATCAGGATATGCCCAAGCACGCGTGTCCAAGCTGCAGTCATCTTTTC 6240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6181 CTGGGATTGGATCAGGATATGCCCAAGCACGCGTGTCCAAGCTGCAGTCATCTTTTC 6240

Qy   6241 TGTTGAGAATGTGTTTGCAAAACTTTACAAAGGACCCCAGAACTGTTCAAATTACTGGAG 6300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6241 TGTTGAGAATGTGTTTGCAAAACTTTACAAAGGACCCCAGAACTGTTCAAATTACTGGAG 6300

Qy   6301 AGGGGCTGTTCCAGTCAACGCGTAGGCGTGTGGGTCGGCTAGACCGGACCCAACTGATTG 6360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6301 AGGGGCTGTTCCAGTCAACGCGTAGGCGTGTGGGTCGGCTAGACCGGACCCAACTGATTG 6360

Qy   6361 GACTAGTCCTTCGTCGTCAATTATGGGCTGCCAAATATGGGACTGTAAATATGAGAAATGGGAGA 6420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6361 GACTAGTCCTTCGTCGTCAATTATGGGCTGCCAAATATGGGACTGTAAATATGAGAAATGGGAGA 6420

Qy   6421 TCACATTTTTGTTACAGCAGTATCCTCCTCCTCCAAATGTCTGTTCACCCCAGTTGCCCCCAAC 6480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6421 TCACATTTTTGTTACAGCAGTATCCTCCTCCTCCAAATGTCTGTTCACCCCAGTTGCCCCCAAC 6480

Qy   6481 CTTGAGAGCTGCAGTCGCCGTGACCGGCGTACACGCTTCAGTGTTATCTGGACCTGATTG 6540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6481 CTTGAGAGCTGCAGTCGCCGTGACCGGCGTACACGCTTCAGTGTTATCTGGACCTGATTG 6540

Qy   6541 AACTCCTTGGACGACATCTGCTTGCTGTTACGGTCCTGACGGTAAGGGTAAAACTGTTAA 6600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6541 AACTCCTTGGACGACATCTGCTTGCTGTTACGGTCCTGACGGTAAGGGTAAAACTGTTAA 6600

Db   6541 AACTCCTTGGACGACATCTGCTTGCTGTTACGGTCCTGACGGTAAGGGTAAAACTGTTAA 6600

Qy   6601 GCTTCCCTTCCCGCGTTGACGGTCACACACTTGGTGTGGCATGCAACTTAATTGCGTGA 6660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6601 GCTTCCCTTCCCGCGTTGAGACTGACTGTAATTCCACAACACACTCCTAGTGATGAAGCCGCAGT 6720

Qy   6661 TGCACTTGAGACAAATGACTGTAATTCCACAACACACTCCTAGTGATGAAGCCGCAGT 6720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6661 TGCACTTGAGACAAATGACTGTAATTCCACAACACACTCCTAGTGATGAAGCCGCAGT 6720

Qy   6721 GTCCGCTCTTGTTTCAAACAGGAGTGCGCGGGTACAAACCAATTGCTTGAGGCAATTTC 6780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6721 GTCCGCTCTTGTTTCAAACAGGAGTGCGCGGGTACAAACCAATTGCTTGAGGCAATTTC 6780

Qy   6781 AGCTGGGCGTTGACACCACCAAACTGCCAGCCCCTCCATCGAAGAGGTAGTGGTAAGAAA 6840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6781 AGCTGGGCGTTGACACCACCAAACTGCCAGCCCCTCCATCGAAGAGGTAGTGGTAAGAAA 6840

Qy   6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGGTTACCTTGCCTCCCCCGGATCCGTCCC 6900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGGTTACCTTGCCTCCCCCGGATCCGTCCC 6900

Qy   6901 AGGAGTGTCATGTCCTGAAAGCTGCAACGAAGTGACCCGTAGAGAGTCCTTCAAACCT 6960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6901 AGGAGTGTCATGTCCTGAAAGCTGCAACGAAGTGACCCGTAGAGAGTCCTTCAAACCT 6960

Qy   6961 CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGCCGATGCCCTGTTGGGAGCGGGGTGA 7020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6961 CCCTCCTTCACCACCTGTTCTACAGTTGGCCATGCCGATGCCCTGTTGGGAGCGGGGTGA 7020

Qy   7021 GTGTAACCCTTCACTGCAATTGGATGTGCAATGACCGGAAACAGGCGGAGGCCCTGATGA 7080
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7021 GTGTAACCCTTCACTGCAATTGGATGTGCAATGACCGGAAACAGGCGGAGGCCCTGATGA 7080

Qy   7081 TTTACCCAGTTACCCTCCCAAAAAGAGGAGTCTCTGAATGGTCAGACGAAAGTTGGTCGAC 7140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7081 TTTACCCAGTTACCCTCCCAAAAAGAGGAGTCTCTGAATGGTCAGACGAAAGTTGGTCGAC 7140

Qy   7141 GGCTACAACCGCTTCCAGCTACGTTACCGGCCCCCCGTACCCTAAGATACGGGGAAAAGA 7200
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7141 GGCTACAACCGCTTCCAGCTACGTTACCGGCCCCCCGTACCCTAAGATACGGGGAAAAGA 7200

Qy   7201 TTCCACTCAGTCAGCCCGCCAAACGGCCAACCGGGGTACAAGGGCCC 7260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7201 TTCCACTCAGTCAGCCCGCCAAACGGCCAACCGGGGTACAAGGGCCC 7260

Qy   7261 TTCGTGGAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7261 TTCGTGGAGCATGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGT 7320

Qy   7321 TCTGTGTCGCAACTCGGGCGATGCGGAGCTCACCAGTGGTGTTCCTCAAACAAGATCATTGGTGTATGT 7380
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7321 TCTGTGTCGCAACTCGGGCGATGCGGAGCTCACCAGTGGTGTTCCTCAAACAAGATCATTGGTGTATGT 7380

Qy   7381 GACTGAGCGCGTGGGGATCGCGGAGCTTAGAAACAAAAAGTCACTATTAATAGACAACCTCT 7440
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7381 GACTGAGCGCGTGGGGATCGCGGAGCTTAGAAACAAAAAGTCACTATTAATAGACAACCTCT 7440

Qy   7441 GTTCCCCCATCATACCACCAAGCAAGTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7441 GTTCCCCCATCATACCACCAAGCAAGTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500

Qy   7501 CGGTGTCATGTGGGGAGCTGGGACTGATGAAGTGAAGCAGCTCACACGCGTCTCTAAGTC 7560
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7501 CGGTGTCATGTGGGGAGCTGGGACTGATGAAGTGAAGCAGCTCACACGCGTCTCTAAGTC 7560

Qy   7561 CCACATCACTGGCCTTCGGGGGACTACTGTTCGTTCGGGGAGCAGCCCCAAGGCTGTTCT 7620
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7561 CCACATCACTGGCCTTCGGGGGACTACTGTTCGTTCGGGGAGCAGCCCCAAGGCTGTTCT 7620

Qy   7621 GGACTTGCAGAAGTGTGTGGAGGCAGGTAAGGGTGAGATACCGGAGTCATTATCGGCAAACTGTGAT 7680
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7621 GGACTTGCAGAAGTGTGTGGAGGCAGGTAAGGGTGAGATACCGGAGTCATTATCGGCAAACTGTGAT 7680

```
Qy  7681  AGTTCCAAAGGAGGAGGAGTCTTCGTGAAGACCCCCCAGAAACCAACAACAAAGAAACCCCCAAG  7740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7681  AGTTCCAAAGGAGGAGGAGTCTTCGTGAAGACCCCCAGAAACCAACAACAAAGAAACCCCCAAG  7740

Qy  7741  GCTTATCTCGTACCCCCACCTTGAAATGAGATGTGTTGAGAGATGTTACTACGGGTCAGGT  7800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7741  GCTTATCTCGTACCCCCACCTTGAAATGAGATGTGTTGAGAGATGTTACTACGGGTCAGGT  7800

Qy  7801  TGCTCCTGACGTAGTTAAAGCTGTCATGGAGATGCGTACGGGTTTGTAGATCCACGTAC  7860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7801  TGCTCCTGACGTAGTTAAAGCTGTCATGGAGATGCGTACGGGTTTGTAGATCCACGTAC  7860

Qy  7861  CCGTGTCAAGCGTCTGTTGTCGATGTGGTCACCCGAGTCGGAGCCACATGCGATAC  7920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7861  CCGTGTCAAGCGTCTGTTGTCGATGTGGTCACCCGAGTCGGAGCCACATGCGATAC  7920

Qy  7921  AGTGTGTTTGACAGTACCATCATCACACCCGAGGATATCATGGTTGGAGACAGACATCTACTC  7980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7921  AGTGTGTTTGACAGTACCATCATCACACCCGAGGATATCATGGTTGGAGACAGACATCTACTC  7980

Qy  7981  AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCATTCACACACCATTGCGAGGCAGTTATA  8040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7981  AGCAGCTAAACTCAGTGACCAACACCGAGCTGGCATTCACACACCATTGCGAGGCAGTTATA  8040

Qy  8041  CGCTGGAGGACCGATGATCGCTTATGATGGCCGAGATCGGATATCGTAGGTGTAGGTC  8100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8041  CGCTGGAGGACCGATGATCGCTTATGATGGCCGAGATCGGATATCGTAGGTGTAGGTC  8100

Qy  8101  TTCCGGGGTCTATACTACCTCAAGTTCCAACAGTTTGACCTGGCTGGCGAAGGTAAATGC  8160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8101  TTCCGGGGTCTATACTACCTCAAGTTCCAACAGTTTGACCTGGCTGGCGAAGGTAAATGC  8160

Qy  8161  TGCAGCCGAACAGGCTGGCGCCGGACAGATGCAGACAAAGCAATGCGTGTCTTTGCTAG  8220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161  TGCAGCCCGAACAGGCTGGCGCCGGACAGATGCAGACAAAGCAATGCGTGTCTTTGCTAG  8220

Qy  8221  CGTAATTTGGAAGAGCGGCCGGAGCCGGAGACGCAGAAGCAATGCGTGTCTTTGCTAG  8280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8221  CGTAATTTGGAAGAGCGGCCGGAGCCGGAGACGCAGAAGCAATGCGTGTCTTTGCTAG  8280

Qy  8281  CTGGATGAAGGTGATGGGTGCACCAAGATTGTGTGCCTCAACCCAAATACAGTTTGGA  8340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281  CTGGATGAAGGTGATGGGTGCACCAAGATTGTGTGCCTCAACCCAAATACAGTTTGGA  8340

Qy  8341  AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA  8400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8341  AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA  8400

Qy  8401  CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTCGCCGAGGGTCTGGG  8460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8401  CTACTTTCTTACAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTCGCCGAGGGTCTGGG  8460

Qy  8461  ATACAACCCCAGTGCGTCGCGTGGATTGGCTGTATCTAATACATCACTACACCCCATGTTTGTGGGT  8520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8461  ATACAACCCCAGTGCGTCGCGTGGATTGGCTGTATCTAATACATCACTACACCCCATGTTTGTGGGT  8520

Qy  8521  TAGCCGTGTGTTGGCTGTCCATTTCATTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA  8580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8521  TAGCCGTGTGTTGGCTGTCCATTTCATTCATGGAGCAGATGCTCTTTGAGGACAAACTTCCCGA  8580

Qy  8581  GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG  8640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8581  GACGGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGCCCAG  8640

Qy  8641  CATCAATTGCTCGGTGTGCACGGGTATTGAGGCTTTCTCGGTGGTGCTCACCCAACGCTGA  8700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8641  CATCAATTGCTCGGTGTGCACGGGTATTGAGGCTTTCTCGGTGGTGCTCACCCAACGCTGA  8700

Qy  8701  GATCCTCAGAGTTTCCCAATCACTAACAGACACTAACAGACCATGCCCCCCTGCGCGAGCCTGGCG  8760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8701  GATCCTCAGAGTTTCCCAATCACTAACAGACACTAACAGACCATGCCCCCCTGCGCGAGCCTGGCG  8760
```

```
Qy  8761  AAAGAAAGCCAGGCGCGGTCCTCCGCCAGCGCCAAGAGGCCGTGGCCGGAGCACGCGAAATT  8820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8761  AAAGAAAGCCAGGCGCGGTCCTCCGCCAGCGCCAAGAGGCCGTGGGCGGAGCACGCGAAATT  8820

Qy  8821  GGCTCGCTTCCTTCTCTGCATGCTACATCTAGACCTCTAGACCTCTACAGATTGGATAAGACGAG  8880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8821  GGCTCGCTTCCTTCTCTGCATGCTACATCTAGACCTCTACAGATTGGATAAGACGAG  8880

Qy  8881  CGTTGGCTCGGTACCACCTTCAATTATTGTGATGTTTACTCCCGGAGGGGGATGTGTT  8940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8881  CGTTGGCTCGGTACCACCTTCAATTATTGTGATGTTTACTCCCGGAGGGGGATGTGTT  8940

Qy  8941  TATTACACCACAGAGAAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTTT  9000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8941  TATTACACCACAGAGAAGATTGCAGAAGTTCCTTGTGAAGTATTTGGCTGTCATTGTTTT  9000

Qy  9001  TGCCCTAGGGCTCATTGCTGTGTTGGATTAGCCATCAGCTTGAACCCCAAATTCAAAATTAA  9060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9001  TGCCCTAGGGCTCATTGCTGTGTTGGATTAGCCATCAGCTGAACCCGAACCCCAAATTCAAAATTAA  9060

Qy  9061  CTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCCACATGTGATCAGAACCGTTCGGGTGAA  9120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9061  CTAACAGTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCCACATGTGATCAGAACCGTTCGGGTGAA  9120

Qy  9121  TTAACGACCCCGCCGATGTGAGTTGGCGAAGCAGTCAGTGGCTCATCCACAAAAACCGTCTCGGGTGG  9180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9121  TTAACGACCCCGCCGATGTGAGTTGGCGAAGCAGTCAGTGGCTCATCCACAAAAACCGTCTCGGGTGG  9180

Qy  9181  GCCATGGTCTGAAGGGGATGACGTCCCTTCTGGCTCATCCGGCTCATCCGGCTCTCGTGTGTGGTGTGACGC  9240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9181  GCCATGGTCTGAAGGGGATGACGTCCCTTCTGGCTCATCCGGCTCATCCACAAAAACCGTCTCGTGTGTGGTGTGACGC  9240

Qy  9241  GTGAGGAGTCCTGCTGCTGTGTGCGAAGCAGTCAGTGGCGTAGTACCAAGGGCTGCACCCCGGTTGTGACCGG  9300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9241  GTGAGGAGTCCTGCTGCTGTGTGCGAAGCAGTCAGTGGCGTAGTACCAAGGGCTGCACCCCGGTTGTGACCGG  9300

Qy  9301  CTCACGACGTATTTGTCCGCTGTGTGCGAGAGCGTGTCGCAGCCGTAGTACACCCCGGTTTTTG  9360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9301  CTCACGACGTATTTGTCCGCTGTGTGCGAGAGCGTGTCGCAGCCGTAGTACACCCCGGTTTTTG  9360

Qy  9361  TTCCAAGCGGAGGCAACCCCGCTTGAATTAAAAACT  9399
          ||||||||||||||||||||||||||||||||||||
Db  9361  TTCCAAGCGGAGGCAACCCCGCTTGAATTAAAAACT  9399
```

```
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  Patentin Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/424,550B
;     FILING DATE:
;     CLASSIFICATION:  435435
;  ATTORNEY/AGENT INFORMATION:
;     NAME:  POREMBSKI, PRISCILLA E.
;     REGISTRATION NUMBER:  33,207
;     REFERENCE/DOCKET NUMBER:  5527.PC.01
;  TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  708-937-6365
;     TELEFAX:  708-938-2623
;
;  INFORMATION FOR SEQ ID NO:  390:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH:  9143 base pairs
;        TYPE:  nucleic acid
;        STRANDEDNESS:  double
;        TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-424-550B-390

Query Match         96.4%;     Score 9059.8;     DB 8;     Length 9143;
Best Local Similarity 99.6%;   Pred. No. 0;
Matches 9096;   Conservative  0;   Mismatches  37;   Indels  4;   Gaps   1;
```

```
Qy       1  ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCTAG   60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1  ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCTAG   60

Qy      61  CAGGGGCTGGGGATTTCCCCTGCCCTCTGCAGAAGGTGAGCCAACCACCTTAGTAT  120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61  CAGGGGCTGGGGATTTCCCCTGCCCTCTGCAGAAGGTGAGCCAACCACCTTAGTAT  120

Qy     121  GTAGGGCGCGGGACTCATGACGCTCGCGTGATGACAAGCGGCAAGCTTGACTTGATGGC  180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121  GTAGGGCGCGGGACTCATGACGCTCGCGTGATGACAAGCGGCAAGCTTGACTTGATGGC  180

Qy     181  CCTGATGGGCGTTCATGGGTTCGGTGTCGGTGGTGGTGGCGCTTAGGCAGCCTCCACGCGCCACCA  240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181  CCTGATGGGCGTTCATGGGTTCGGTGTCGGTGGTGGTGGCGCTTAGGCAGCCTCCACGCGCCACCA  240

Qy     241  CCTCCCAGATAGAGCGCGGCGCGGCGCACTGTAGGGAGACCCGGGGACCACCAGGACG  300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241  CCTCCCAGATAGAGCGCGGCGCGGCGCACTGTAGGGAGACCCGGGGACCACCAGGACG  300

Qy     301  CAGAGACCTCTTTTGAGTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTGT  360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301  CAGAGACCTCTTTTGAGTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACCTATATGTGT  360

Qy     361  TGGGGATGGTTGGGGTTAGCCATCCATCACCGCTACTACCGTCCGCCTTGCCTTGGCCGTCCTTGCCTGAGGGGGAT  420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361  TGGGGATGGTTGGGGTTAGCCATCCATCACCGCTACTACCGTCCGCCTTGCCTTGGCCGTCCTTGCCTGAGGGGGAT  420

Qy     421  CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTTCTACTCAAACAGTCCTGTACC  480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421  CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTTCTACTCAAACAGTCCTGTACC  480

Qy     481  TGCGCCCAGAACGCGCAAGAACAAGCAGAGCAAAGCGCAAAGTCCAGGCGCGATGCTCGCCGCCTTTAAAAC  540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481  TGCGCCCAGAACGCGCAAGAACAAGCAGAGCAAAGCGCAAAGTCCAGGCGCGATGCTCGCCGCCTTTAAAAC  540

Qy     541  ATCTGTTGAAAGGGGACCAACGACAAGCTTGCCAGACATTGGCTCAGGCTGCTTTGCCAGC  600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541  ATCTGTTGAAAGGGGACCAACGACAAGCTTGCCAGACATTGGCTCAGGCTGCTTTGCCAGC  600

Qy     601  TTACAAAATTGCTGGTATCCATGATGGCCCATAAGTCTCCGCAATCTTCGGAATCTTCTTGGA  660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601  TTACAAAATTGCTGGTATCCATGATGGCCCATAAGTCTCCGCAATCTTCGGAATCTTCTTGGA  660

Qy     661  TCATGGTTGGGGACGGACGCCAAGACCCTCGCCATAAGTCTCGCCAATCTTCGGAATCTTCTTCTTGGA  720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661  TCATGGTTGGGGACGGACGCCAAGACCCTCGCCATAAGTCTCGCCAATCTTCGGAATCTTCTTCTTGGA  720
```

```
Qy     721  TTACCCTTGGGGTGGGATTGGTGATGTTGATGTTACAACTCACACACCTCTAGTAGGCCCGCTGGT  780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721  TTACCCTTGGGGTGGGATTGGTGATGTTGATGTTACAACTCACACACCTCTAGTAGGCCCGCTGGT  780

Qy     781  GGCAGGAGGAGCGGTCGTTGCGACCAGTCTGCTTGCTGCCAGATAGTACGCTTGCTGGAGGATGGAGTCAA  840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781  GGCAGGAGGAGCGGTCGTTGCGACCAGTCTGCTTGCTGCCAGATAGTACGCTTGCTGGAGGATGGAGTCAA  840

Qy     841  CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTGTGTGATGTCTGCTATCTTTGGCCTG  900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841  CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTGTGTGATGTCTGCTATCTTTGGCCTG  900

Qy     901  TCCCTGTAGTGGAGGCGCGGCGCGGGCGGGTCACTGACCCAGACACAAATACCACACATCCTGACCAATTG  960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901  TCCCTGTAGTGGAGGCGCGGCGCGGGCGGGTCACTGACCCAGACACAAATACCACACATCCTGACCAATTG  960

Qy     961  CTGCCAGCGTAATCAGTTATCTATTGTCTCCTCCTTCCACTTGCCTACACGAGCCTGGTTG  1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961  CTGCCAGCGTAATCAGTTATCTATTGTCTCCTCCTTCCACTTGCCTACACGAGCCTGGTTG  1020

Qy    1021  TGTGATCGTGCGGAGGAGTGCTGGGTTCGCGCCAATCCGTACAATCTCACACCCTTCCAA  1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021  TGTGATCGTGCGGAGGAGTGCTGGGTTCGCGCCAATCCGTACAATCTCACACCCTTCCAA  1080

Qy    1081  TTGGACTGGCACGGACTCCTTCTTGGCTGACCACCATGATTTTGTTATGGGGCGCTCTTGT  1140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081  TTGGACTGGCACGGACTCCTTCTTGGCTGACCACCATGATTTTGTTATGGGGCGCTCTTGT  1140

Qy    1141  GACCTGTGACGCCCTTGACATTGGTGAGTTGTGTGGCGTGGTATTAGTCGGTGACTG  1200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141  GACCTGTGACGCCCTTGACATTGGTGAGTTGTGTGGCGTGGTATTAGTCGGTGACTG  1200

Qy    1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGA  1260
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201  GCTTGTCAGGCACTGGCTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGA  1260

Qy    1261  AGTGCCACTGGAATAGATCCTGGGTTCCTAGGTTTATCGGGGTGGATGGCCGGCAAGGT  1320
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261  AGTGCCACTGGAATAGATCCTGGGTTCCTAGGTTTATCGGGGTGGATGGCCGGCAAGGT  1320

Qy    1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT  1380
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321  CGAGGCTGTCATCTTCTTGACCAAACTGGCTTCACAAGTACCATACGCTATTGCGACTAT  1380

Qy    1381  GTTTAGCAGTGTACACTACCTGGCGCGTTGGCGCTTTACATAGAAGCGACCTCTCGGGGCAA  1440
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381  GTTTAGCAGTGTACACTACCTGGCGCGTTGGCGCTTTACATAGAAGCGACCTCTCGGGGCAA  1440

Qy    1441  GTGGTATCAGTTGCTCCTAGCGTTATGCTTACATAGAAGCGACCTCTGATACCATGTCC  1500
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441  GTGGTATCAGTTGCTCCTAGCGTTATGCTTACATAGAAGCGACCTCTGATACCATGTCC  1500

Qy    1501  CAGGGTGCCCACTGGGATGCTCAATAGCTGAGTTTGCTCGCCCTTGATGATACCAAGTGGAC  1560
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501  CAGGGTGCCCACTGGGATGCTCAATAGCTGAGTTTGCTCGCCCTTGATGATACCAAGTGGAC  1560

Qy    1561  TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTGTCAGAAGTGCCAAGTGGAC  1620
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561  TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTGTCAGAAGTGCCAAGTGGAC  1620

Qy    1621  CAGGCTTATCATCTAGATATAACAACTCCATATCTTGGTACCCTATACAATCCCTGG  1680
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621  CAGGCTTATCATCTAGATATAACAACTCCATATCTTGGTACCCTATACAATCCCTGG  1680

Qy    1681  TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCGTATTCCGCAA  1740
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681  TGCGAGGGGATGTATGGTTAAATTCAAAAATAACACATGGGGTTGCTGCGTATTCGCAA  1740

Qy    1741  TGTGCCATCGTACTGCACTATGGGACACTGATGCGAGTGTGGAACGACACTCGCAACACTTA  1800
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741  TGTGCCATCGTACTGCACTATGGGACACTGATGCGAGTGTGGAACGACACTCGCAACACTTA  1800
```

Qy   1801 CGAAGCAATGCGGTGTAACACCATGGCTAACAACCGCATGGCACAACGGCTCAGCCCTTGAA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CGAAGCAATGCGGTGTAACACCATGGCTAACAACCGCATGGCACAACGGCTCAGCCCTTGAA 1860

Qy   1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAATGTTAAACCTCATAATTGGATGTC 1920

Qy   1921 AGGCCATTTGTATTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA 1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AGGCCATTTGTATTTGAGGGATCAGATACCCTATAGTTTACTTTTATGACCCTGTGAA 1980

Qy   1981 TTCCACTCTCCTACCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCTGTGTACG 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TTCCACTCTCCTACCACCGGAGAGGTGGGCTAGGTTGCCCGGTACCCCACCTGTGTACG 2040

Qy   2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTACAGTGATGTGAAAGACCTAGCCACAGG 2100

Qy   2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGTCTCTATATTCCGCCACGGGTGC 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGTCTCTATATTCCGCCACGGGTGC 2160

Qy   2161 TTTGTCTTCTTACGGGAGTTACCACCAAGGCCGTGGTGTGCTAATTCGTGTGGGGTGTGG 2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TTTGTCTTCTTACGGGAGTTACCACCAAGGCCGTGGTGTGCTAATTCGTGTGGGGTGTGG 2220

Qy   2221 CAGCAAGTATCTTATTTTAGCCTCTGTTACTTGTCCCTTTGTTTGGGCGCCGTTC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CAGCAAGTATCTTATTTTAGCCTCTGTTACTTGTCCCTTTGTTTGGGCGCCGTTC 2280

Qy   2281 TGGTTACCCTTTGCGTCCTGTGCTCCATCCCAGTCGTATCTTCCAAGCTGGCTGGGATGT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 TGGTTACCCTTTGCGTCCTGTGCTCCATCCCAGTCGTATCTTCCAAGCTGGCTGGGATGT 2340

Qy   2341 TTTGTCTAAAGCTCAAGTAGTTCCTTTGCTTTGATTTTCTTCATCGTTGCTATCTCCG 2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TTTGTCTAAAGCTCAAGTAGTTCCTTTGCTTTGATTTTCTTCATCGTTGCTATCTCCG 2400

Qy   2401 CTGCAGGCTACGTTATGCTGCCCTTCTTGCATGGCTGTACGCCTGCTTCAGC 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 CTGCAGGCTACGTTATGCTGCCCTTCTTGCATGGCTGTACGCCTGCTTCAGC 2460

Qy   2461 AACTTTCTTGTTGCAGCGATGATAATTGGAGGGCTGACAATACCACCTGTTAGTTGTCAT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 AACTTTCTTGTTGCAGCGATGATAATTGGAGGGCTGACAATACCACCTGTTAGTTGTCAT 2520

Qy   2521 AGTGGCGAGGGTTAGTTTGTGGGGCCGGCCGCTAACCGGTCACCGCATAGCTCTGCTTGT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 AGTGGCGAGGGTTAGTTTGTGGGGCCGGCCGCTAACCGGTCACCGCATAGCTCTGCTTGT 2580

Qy   2581 AGTTCCTTGGCCTCTCGGTAGCGCTTTTAACCCTCTTGCATTTGGTTACGCCTGCTTCAGC 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 AGTTCCTTGGCCTCTCGGTAGCGCTTTTAACCCTCTTGCATTTGGTTACGCCTGCTTCAGC 2640

Qy   2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTTAGTAGCATTAGTTGTCAT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTTAGTAGCATTAGTTGTCAT 2700

Qy   2701 GTCTCGTTTGGCTTCTTGCTCACTTGTACCTCGCTGTGCTTTAGTTAAACTCCTATCT 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GTCTCGTTTGGCTTCTTGCTCACTTGTACCTCGCTGTGCTTTAGTTAAACTCCTATCT 2760

Qy   2761 TTTGGCAACGTTGGGAGAATTGGTTTTGGAACGTTACACGTTGTTTGGAACGTTACAC... 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 TTTGGCAACGTTGGGAGAATTGGTTTTGGAACGTTACACGTTGTTTGGAACGTTACAC... 2820

Qy   2821 TGTGTGCTGGTTGTTGTTCCCCGGTGCGACATATGACCGCGCTGGTGACTTTCTGTGTGTCA 2880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 TGTGTGCTGGTTGTTGTTCCCCGGTGCGACATATGACCGCGCTGGTGACTTTCTGTGTGTCA 2880

Qy   2881 CGTAGCGCTTCTATGTTTAACATCCAGTGCAGCAATCGTTCTTTGGGACTGACTCTAGGGT 2940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 CGTAGCGCTTCTATGTTTAACATCCAGTGCAGCAATCGTTCTTTGGGACTGACTCTAGGGT 2940

Qy   2941 TAGGGCCCATAGAATGTTGGTGCGTCTCGGAAAGTGTCATGCGTATTCTCATTATGT 3000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 TAGGGCCCATAGAATGTTGGTGCGTCTCGGAAAGTGTCATGCGTATTCTCATTATGT 3000

Qy   3001 TCTTAAGTTTTCCTCTAGTGTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA 3060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 TCTTAAGTTTTCCTCTAGTGTTGGTGAGAATGGTGTGTTTTCTATAAGCACTTGCA 3060

Qy   3061 TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCAAGAGCCATTTTCCC 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 TGGTGATGTCTTGCCTAATGATTTGCCTCGAAACTACCATTGCAAGAGCCATTTTCCC 3120

Qy   3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGGAAGACGCTTGGGCGTTGTGGGGACAC 3180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 TTTTGAAGGCAAGGCAAGGGTCTATAGGAATGAAGGAAGGAAGACGCTTGGGCGTTGTGGGGACAC 3180

Qy   3181 GGTTGATGTTGCCCGTGTTGCCCGTCTCGGCGCTCGGCGCACCCTTGTTTCGCAGGGGTTGGCTAT 3240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 GGTTGATGTTGCCCGTGTTGCCCGTCTCGGCGCTCGGCGCACCCTTGTTTCGCAGGGGTTAGCTAT 3240

Qy   3241 GCCGCCAGATGGGTGGGCCATTACCGCACCTTTTACGCTGCGCAGTGTCTCTCTGAACGTGG 3300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 GCCGCCAGATGGGTGGGCCATTACCGCACCTTTTACGCTGCGCAGTGTCTCTCTGAACGTGG 3300

Qy   3301 CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTTGGGAAC 3360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 CACGCTGTCAGCGATGGCAGTGGTCATGACTGGTATAGACCCCGAACTTGGACTTGGGAAC 3360

Qy   3361 TATCTTCAGATTAGGATCTTCGGGCCACTAGCTACACATGGGATTGTTGTTGTGACAACGTGTT 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 TATCTTCAGATTAGGATCTTCGGGCCACTAGCTACACATGGGATTGTTGTTGTGACAACGTGTT 3420

Qy   3421 GTATACTGCTCACCATGGCAGCAAGGGGCGCCGGTTGGCTCATCCCGGTTGGCTCCATCACA 3480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 GTATACTGCTCACCATGGCAGCAAGGGGCGCCGGTTGGCTCATCCCGGTTGGCTCCATCACA 3480

Qy   3481 CCCAATAACCGTTGACGCGGCTAATGACGCGGCATCTATCAACCACCATGTGGAGCTGG 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 CCCAATAACCGTTGACGCGGCTAATGACGCGGCATCTATCAACCACCATGTGGAGCTGG 3540

Qy   3541 GTCCCTTACTCGGTGCTCTTGCGGGGAGACCAAGGGTATCTGGGGTAACACGACTGGGGTC 3600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 GTCCCTTACTCGGTGCTCTTGCGGGGAGACCAAGGGTATCTGGGGTAACACGACTGGGGTC 3600

Qy   3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGGTGCGGGGCCCTTCCCAT 3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 ATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGGTGCGGGGCCCTTCCCAT 3660

Qy   3661 GGCTGTTGCCAAGGGTTCTTCAGGTGCCCGATTCGTGTGCTCGTGTGCATGGCAGCGGCAAGTCAAC 3720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 GGCGTTGCCAAGGGTTCTTCAGGTGCCCGATTCGTGTGCTCGTGTGCATGGCAGCGGCAAGTCAAC 3720

Qy   3721 GATGTTCACCGGCTGCTAGAAATTCTGGGCGTTCAGTCGTTCAGTCAGCGGTTAGGCCGTT 3780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 GATGTTCACCGGCTGCTAGAAATTCTGGGCGTTCAGTCGTTCAGTCAGCGGTTAGGCCGTT 3780

Qy   3781 GGTGTGTCGTGCTGGATACCATCCCAGTACACAGCACACATGCCACTCTTGATACAAAACCTAC 3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 GGTGTGTCGTGCTGGATACCATCCCAGTACACAGCACACATGCCACTCTTGATACAAAACCTAC 3840

Qy   3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC 3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC 3900

Qy   3901 CAAATTACCACTTTCTTACCACTTTCTTACATGCAGGAGAAGTATGAGGTCCTGGTCCTCAGTGT 3960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 CAAATTACCACTTTCTTACCACTTTCTTACATGCAGGAGAAGTATGAGGTCCTGGTCCTCAGTGT 3960

Qy   3961 GGCTACAACAGCATCAATGCCAAAGTACAATGCCGACGTACGGCGTACGGGGT 4020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db   3961 GGCTACAACAGCATCAATGCACGCGACGTACGGCGTGAATCCAAATTG 4020

Qy   4021 CTATTTTAATGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGCGCAT 4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 CTATTTTAATGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGCGCAT 4080

Qy   4081 GTACCTGACCGGAGCATGTTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGC 4140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 GTACCTGACCGGAGCATGTTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGC 4140

Qy   4141 TACCGATGCAACCACCACCGTGTTGGGCATTGGGAAAGGTCCTAACCGAAGCTTCCATCC 4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 TACCGATGCAACCACCACCGTGTTGGGCATTGGGAAAGGTCCTAACCGAAGCTTCCATCC 4200

Qy   4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCTGGAGTAATCCCTACACCACATGC 4260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCTGGAGTAATCCCTACACCACATGC 4260

Qy   4261 CAACATAACTGAGATTCAATTAACCAGTTAACCGATGAAGGCACTATCCCCTTTCATGAAAAAAGAT 4320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 CAACATAACTGAGATTCAATTAACCAGTTAACCGATGAAGGCACTATCCCCTTTCATGAAAAAAGAT 4320

Qy   4321 TAAGGAGAAAATCTGAAGAAAGGGAGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321 TAAGGAGAAAATCTGAAGAAAGGGAGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4380

Qy   4381 TGATGAGCTTGCTAACGAGTTAGCTGAAGGAATAACAGCTGTCTCTTACTACTATAGGGG 4440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 TGATGAGCTTGCTAACGAGTTAGCTGAAGGAATAACAGCTGTCTCTTACTACTATAGGGG 4440

Qy   4441 ATGTGACATCTCAAAAATCCCTGAGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 ATGTGACATCTCAAAAATCCCTGAGGGCGACTGTGTAGTAGTTGCCACTGATGCCTTGTG 4500

Qy   4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTGTATGACTGCAGCCTCATGGTAGAAGGCAC 4560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 TACAGGGTACACTGGTGACTTTGATTCCGTGTGTATGACTGCAGCCTCATGGTAGAAGGCAC 4560

Qy   4561 ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTGTTCGTGTGCGCGGGTTTCAGC 4620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 ATGCCATGTTGACCTTGACCCTACTTTCACCATGGGTGTTCGTGTGCGCGGGTTTCAGC 4620

Qy   4621 AATAGTTAAAGGCCAGCCGTAGGGGCCGCACAGGGCCCGTGGGGTATGTTCCGGGAGAGCCGTGGCATATACTACTA 4680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4621 AATAGTTAAAGGCCAGCCGTAGGGGCCGCACAGGGCCCGTGGGGTATGTTCCGGGAGAGCCGTGGCATATACTACTA 4680

Qy   4741 TGTAGACGGGAGTTGTACCCCTTCGGGTATGTTTGTCATCAACATTGTTGAAGCCTT 4800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 TGTAGACGGGAGTTGTACCCCTTCGGGTATGTTTGTCATCAACATTGTTGAAGCCTT 4800

Qy   4801 CGACGCAGCCAAGGCGCTGTCCTGGCCCAAACTGGGGTATGTTGTCATCAACATATTCTGACACT 4860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 CGACGCAGCCAAGGCGCTGTCCTGGCCCAAACTGGGGTATGTTGTCATCAACATATTCTGACACT 4860

Qy   4861 CTTTTCTATGGTCAACCCGAACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA 4920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861 CTTTTCTATGGTCAACCCGAACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA 4920

Qy   4921 TTATGTTTGTTGTGACTGCAGCGCCAACTACAACTGTCATCAGTATGCTATGCTGCTCC 4980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 TTATGTTTGTTGTGACTGCAGCGCCAACTACAACTGTCATCAGTATGCTATGCTGCTCC 4980

Qy   4981 CAATGACGACCACCGTGGCCAGGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCGTGTG 5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 CAATGACGACCACCGTGGCCAGGGAGCCCGGCTTGGGAAAAAACCTTGTGGGGTTCGTGTG 5040

Qy   5041 GCGCTTGGACGGCGCTGACCGCCTGTCCTGGCCCCAGAGCCCAGCGCCGAGTGACCAGATACCA 5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041 GCGCTTGGACGGCGCTGACCGCCTGTCCTGGCCCCAGAGCCCAGCGCCGAGTGACCAGATACCA 5100
```

```
Qy   5101 AATGTGCTTCACTGAGTCACTGAGTCAATACTTCTGGGACAGCCCACTCGCTGTCGTGGCGTTGTGGCGTTGGCAGT 5160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101 AATGTGCTTCACTGAGTCACTGAGTCAATACTTCTGGGACAGCCCACTCGCTGTCGTGGCGTTGTGGCGTTGGCAGT 5160

Qy   5161 GGCTATGCTTATCTAGCTAGCCATTGACACTTTGTGCGCCACTTTGTGCGGCGGCGGAGAAGAAGAAATCGT 5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161 GGCTATGCTTATCTAGCTAGCCATTGACACTTTGTGCGCCACTTTGTGCGGCGGCGGAGAAGAAGAAATCGT 5220

Qy   5221 TATTACATCAGTCCTACCGGTGCTACTGTCGCCCCAGTGGTGACGGTGACGGTGACGAAGAAGAAATCGT 5280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221 TATTACATCAGTCCTACCGGTGCTACTGTCGCCCCAGTGGTGACGGTGACGGTGACGAAGAAGAAATCGT 5280

Qy   5281 GGAGGAGTGTGCATCATTCATTCCCTGGAGGCCATGGTTGCTGCAATTGACAAGCTGAA 5340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5281 GGAGGAGTGTGCATCATTCATTCCCTGGAGGCCATGGTTGCTGCAATCGATAAGCTGAA 5340

Qy   5341 GAGTACAATCACCACAACTAGTCCTTCACATTGGAAACCCCTTGAAAAACTTAACAC 5400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5341 GAGTACAATAACCACAACTAGTCCTTCACATTGGAAAACCCTTGAAAAACTTAACAC 5400

Qy   5401 CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTACATCCTTGTTGCTTTCATTGGGTATTACTAC 5460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5401 CTTTCTTGGGCCTCATGCAGCTACAATCCTTGCTACATCCTTGTTGCTTTCATTGGGTATTACTAC 5460

Qy   5461 CACTTTACCTGACAATCCCTTGCTTGTTTGCTTTCATTGGGGTATTACTAC 5520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5461 CACTTTACCTGACAATCCCTTGCTTGTTTGCTTTCATTGGGGTATTACTAC 5520

Qy   5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTGGAGGCGCAATTGCGTCCAA 5580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5521 CCCACTACCTCACAAGATCAAAATGTTCCTGTCATTATTGGAGGCGCAATTGCGTCCAA 5580

Qy   5581 GCTTACAGACGTGTAGAGGCGCACTGGCGTTCATGATGGCCGGGGCTGCGGGGAACAGCTCT 5640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5581 GCTTACAGACGTGTAGAGGCGCACTGGCGTTCATGATGGCCGGGGCTGCGGGGAACAGCTCT 5640

Qy   5641 TGGTACATGGACATGGCGTGGGTTTGTCTTGTCATGCTCTAGGCGGCAGCTCGGCCCTC 5700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5641 TGGTACATGGACATCGGTGGGTTTGTCTTGTCATGCTCTAGGCGGCAGCTCGGCCCTC 5700

Qy   5701 ATCCACTGCTTGTTGACACTTTAAATGCTTTGATGGGTCATCCGGCCCAATGGATCAGCT 5760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5701 ATCCACTGCTTGTTGACACTTTAAATGCTTTGATGGGTCATCCGGCCCAATGGATCAGC 5760

Qy   5761 TGCTGGTTAGTCTACTCCCGTTGACAACAGAGGGCCCAGGAGTTGTGGGCGTCTTGTCAGC 5820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5761 TGCTGGTTAGTCTACTCCCGTTGACAACAGAGGGCCCAGGAGTTGTGGGCGTCTTGTCAGC 5820

Qy   5821 TTGTGCAATGTTGCTTTGACAACAGAGGGCCCAGATCACTGGGCGTCTTGTCCTTAC 5880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5821 TTGTGCAATGTTGCTTTGACAACAGAGGGCCCAGATCACTGGGCGTCTTGTCCTTAC 5880

Qy   5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGTCACATCCG 5940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5881 TATGCTTGCTAGGAGCAACACTGTATGTAATGAGTACTTTATTGCCACTCGTGTCACATCCG 5940

Qy   5941 CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCCGGAGTGTCATATCAGCTTGCAT 6000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5941 CAGGAAGATACTGGGCATTCTGGAGGCATCTACCCCCCGGAGTGTCATATCAGCTTGCAT 6000

Qy   6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGGCCTCATTGCTTGGGGTCTAGAGAT 6060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6001 CCGTTGGCTCCACACCCCGACGGAGGATGATTGCGGGCCTCATTGCTTGGGGTCTAGAGAT 6060

Qy   6061 TTGGCAGTATGTGTGCAATTTCTTTGTGCTTTTAATGTCCTTTAAAGCTGGAGTTCA 6120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6061 TTGGCAGTATGTGTGCAATTTCTTTGTGCTTTTAATGTCCTTTAAAGCTGGAGTTCA 6120

Qy   6121 GAGCATGGTTAACATTCCTGGTTGTCCTTTCACAGCTGCCACAGCTGCCACAAGGGCCC 6180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6121 GAGCATGGTTAACATTCCTGGTTGTCCTTTCACAGCTGCCACAGCTGCCACAAGGGCCC 6180
```

Qy   6181 CTGGATTGGATCAGGTATGCTCCAAGCACGCTGTCCATGCGGTGCTGAACTCATCTTTTC 6240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6181 CTGGATTGGATCAGGTATGCTCCAAGCACGCTGTCCATGCGGTGCTGAACTCATCTTTTC 6240

Qy   6241 TGTTGAGAAATGGTTTGCAAAACTTTACAAAGACCCAGAACTTGTTCAAATTACTGGAG 6300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6241 TGTTGAGAAATGGTTTGCAAAACTTTACAAAGACCCAGAACTTGTTCAAATTACTGGAG 6300

Qy   6301 AGGGGCTGTTCCAGTCAACGCTAGGCTGTGGGTCGGCTAGACCGGACCCAACTGATTG 6360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6301 AGGGGCTGTTCCAGTCAACGCTAGGCTGTGGGTCGGCTAGACCGGACCCAACTGATTG 6360

Qy   6361 GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA 6420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6361 GACTAGTCTTGTCGTCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAAATGGGAGA 6420

Qy   6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCCAAC 6480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6421 TCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCCAAC 6480

Qy   6481 CTTGAGAGAGCTGCAGTGGCCGTGCTGAGACGGCGTACGGTTATCGGAGGTAAAACTGTTAA 6540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6481 CTTGAGAGAGCTGCAGTGGCCGTGCTGAGACGGCGTACGGTTATCGGAGGTAAAACTGTTAA 6540

Qy   6541 AACTCCTTGGACACATCGCTTGCTGTCGTTGACGGTCACACACCTGTGTCGCGCATGCAACTGTGA 6600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6541 AACTCCTTGGACACATCGCTTGCTGTCGTTGACGGTCACACACCTGTGTCGCGCATGCAACTGTGA 6600

Qy   6601 GCTTCCCTTCGCCGGTTGACGGTCACCAGCCCAAACTGCCACACACTGGTGTGCGGCATGCCAACTGCCAGC 6660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6601 GCTTCCCTTCGCCGGTTGACGGTCACCAGCCCAAACTGCCACACACTGGTGTGCGGCATGCCAACTGCCAGC 6660

Qy   6661 TGCACTTGAGACAAATGACTGTAATTCCACAAAACAACTCCAGTGATGAGCCGCAGT 6720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6661 TGCACTTGAGACAAATGACTGTAATTCCATAAAACAACTCCAGTGATGAGCCGCAGT 6720

Qy   6721 GTCCGGCTTCTTGTTTTCAAACAGGAGGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC 6780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6721 GTCCGGCTTCTTGTTTTCAAACAGGAGGTTGCGGCGTACAAACCAATTGCTTGAGGCAATTTC 6780

Qy   6781 AGCTGGGCGTTGACACCACCAAACTGCCAGCCCCATCGAAGAGGTAGTGTTGGGAGAAA 6840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6781 AGCTGGGCGTTGACACCACCAAACTGCCAGCCCCATCGAAGAGGTAGTGTTGGGAGAAA 6840

Qy   6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTGGCCTCCCCCTCCGAGATCCGTCCC 6900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6841 GCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTGGCCTCCCCCTCCGAGATCCGTCCC 6900

Qy   6901 AGGAGTGTCATGTCCTGAAAGCTGACGAGTGGTCTCGATGGTCCTTCAAACCT 6960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6901 AGGAGTGTCATGTCCTGAAAGCTGACGAGTGGTCTCGATGGTCCTTCAAACCT 6960

Qy   6961 CCCTCCTTCACCACCTGTTCTCACAGTTGGCCATGCCCGTGTTGGGAGCGGGTGA 7020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6961 CCCTCCTTCACCACCTGTTCTCACAGTTGGCCATGCCCGTGTTGGGAGCGGGTGA 7020

Qy   7021 GTGTAACCCTTTCACTGCAATTGCACTGGATGTGCAATGACCCAAACAGGCCCTGATGA 7080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7021 GTGTAACCCTTTCACTGCAATTGCACTGGATGTGCAATGACCCAAACAGGCCCTGATGA 7080

Qy   7081 TTTACCCAGTTACCCTCCCAAAAGGAGGTCTCGAATGGTCAGACGAAGTTGGTCGAC 7140
           |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   7081 TTTACCCAGTTACCCTCCCAAAAGGAGGTCTCGAATGGTCGAAGTTGGTCAAC 7140

Qy   7141 GGCTACAACCGCTTCCAGCTACGTTACTGGCCCCCGTACCCTAAGATACGGGGAAGGA 7200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7141 GACTACAACCGCTTCCAGCTACGTTACTGGCCCCCGTACCCTAAGATACGGGGAAGGA 7200

Qy   7201 TTCCACTCAGTCAGCGGGCCCAAACGGCTACACCGGTGGGAAAAGAGTGGGAAAGAGTGAGTT 7260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7201 TTCCACTCAGTCAGCGGGCCCAAACGGCTACACCGGTGGGAAAAGAGTGGGAAAGAGTGAGTT 7260

Qy   7261 TTCGTGCGAGCATGAGCTACACTTGGACCGACGTACACTTGGTGTTTCCTCAAACAAAGATCGCTTCTAAAGT 7320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7261 TTCTGTCTGCAACTCGGGCCCATCAGTTGGTGTTTCCTCAAACAAAGATCGCTTCTAAAGT 7320

Qy   7321 TCTGTCTGCAACTCGGGCCCATCATTACTAGTGTGTTCCTCAAACAAAGATCGCTTCTAAAGT 7380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7321 TCTGTCTGCAACTCGGGCCCATCATTACTAGTGTGTTCCTCAAACAAAGATCATTCATTGGTGTATGT 7380

Qy   7381 GACTGAGCCGCGGCGGATGCGGAGCTTAGAAACATCACTATTAATAGACAACCTCT 7440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7381 GACTGAGCCGCGGCGGATGCGGAGCTTAGAAACATCACTATTAATAGACAACCTCT 7440

Qy   7441 GTTCCCCCATCATACCACCAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7441 GTTCCCCCATCATACCACCAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGT 7500

Qy   7501 CGGTGTCATGTGGGACACTATGAAGTAGCAGCTCACACGCCCTCACGCCCTCCTAAGTCTGCTAAGTC 7560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7501 CGGTGTCATGTGGGACACTATGAAGTAGCAGCTCACACGCCCTCACGCCCTCCTAAGTCTGCTAAGTC 7560

Qy   7561 CCACATCACTGGCCTTCGCGGGCACTGATGTTCGTTCTGAGAACCGCAGCCCCGGAAGGCTGTTCT 7620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7561 CCACATCACTGGCCTTCGCGGGCACTGATGTTCGTTCTGAGAACCGCAGCCCCGGAAGGCTGTTCT 7620

Qy   7621 GGACTTGCAGAAGTGTGTCGAGCGAGGTCACTGATGTTCGTGGAGACCGCGAGTACCGCAGACCCCCAAG 7680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7621 GGACTTGCAGAAGTGTGTCGAGCGAGGTCACTGATGTTCGTGGAGACCGCGAGTACCGCAGACCCCCAAG 7680

Qy   7681 AGTTCCAAAGAGGAGGAGGTCTTCGTGAAGACCCCAGAAACCAACAAAGAAAACCCCCAAG 7740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7681 AGTTCCAAAGAGGAGGAGGTCTTCGTGAAGACGAGTACCGCAGAAACCAACAAAGAAAACCCCCAAG 7740

Qy   7741 TGCACTTGAGACAAATGACTGTAATTCCACAAAACAACTCCAGTGATGAAATGAGATGTGTTGAGAAGATGTACTGGTCAGGT 7800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7741 GCTTATCTCGTACCCCCACCTTGAAATGAGATGTGTTGAGAAGATGTGTTGAGAAGATGTACTGGTCAGGT 7800

Qy   7801 TGCCTCCTGACGTAGTAGTTCATCGGCGTAAAGCGTCATGGGAGAGTCGTACGGGTTTGTGATCCACCGTAC 7860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7801 TGTCCTCGACGTAGTAGTTCATCGGCGTAAAGCGTCATGGGAGAGTCGTACGGGTTTGTCGACCCACCGTAC 7860

Qy   7861 CCGTGTCAAGGCGTCTGTTGTGTCGATGTGGTCACCCGAGCGCACATGCGGATAC 7920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7861 CCGTGTCAAGGCGTCTGTTGTGTCGATGTGGTCACCCGAGCGCACATGCGGATAC 7920

Qy   7921 AGTGTTTTGACAGTACCATCACACCCGAGGATATCATGGTGAGACAGACATCTACTC 7980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7921 AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTGGAGACAGACATCTACTC 7980

Qy   7981 AGCAGCTAAACTCAGTGACCAACACCGAGGCCATTGCCGGCAGCGTCAGCAGTTATA 8040
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7981 AGCAGCTAAACTCAGTGACCAACACCGAGGCCATTGCCGGCAGCGTCAGCAGTTATA 8040

Qy   8041 CGCTTGGAGGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTAGGTC 8100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8041 CGCTTGGAGGACCGATGATCGCTTATGATGGCCGAGAGATCGGATATCGTAGGTGTAGGTC 8100

Qy   8101 TTCCGGCGTCTATACTACCTCAAGTTGACCAGTTTGACCTGCTGGCTGAAGGTAAATGC 8160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8101 TTCCGGCGTCTATACTACCTCAAGTTGACCAGTTTGACCTGCTGGCTGAAGGTAAATGC 8160

Qy   8161 TGCAGCCGAACAGGCCTGCAGGCTGGCATGAGAAGAACCCTCGCTTCCTTCTATTGCGGGCGATGATTGCAC 8220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8161 TGCAGCCGAACAGGCCTGCAGGCTGGCATGAGAAGAACCCTCGCTTCCTTCTATTGCGGGCGATGATTGCAC 8220

Qy   8221 CGTAATTTGGAAGAGCCGGAGGCGCCGGATCGAGACAGACAAGCAACAAGCAATGCGTGTCTTTGCTAG 8280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8221 CGTAATTTGGAGAGCCGGAGGCGCCGGATCGAGACAGACAAGCAACAAGCAATGCGTGTCTTTGCTAG 8280

Qy   8281 CTGGATGAAGGTGATGGGGTGCACCACCAAATACAGTTTGGA 8340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8281 CTGGATGAAGGTGATGGGGTGCACCACCAAGTACAGTTTGGA 8340

Qy   8341 AGAATTAACATCATGCTCATCAAATGTTACCTCCTGGAAATTACCGAAGCCTTA 8400
           |||||||||||||||||||||||

CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES   D377/AP6D
STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,550B
FILING DATE:
CLASSIFICATION: 435435
ATTORNEY/AGENT INFORMATION:
NAME: POREMBSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5527.PC.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 393:
SEQUENCE CHARACTERISTICS:
LENGTH: 9143 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..445
FEATURE:
NAME/KEY: CDS
LOCATION: 446..9037
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 9038..9143
US-08-424-550B-393

Query Match   96.4%;   Score 9059.8;   DB 8;   Length 9143;
Best Local Similarity 99.6%;   Pred. No. 0;
Matches 9096;   Conservative 0;   Mismatches 37;   Indels 4;   Gaps 1;

Qy   1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACCACAAACACTCCAGTTGTTACACTCCGCTAGGAATGCTCCTGGAGCACCCCCCTAG 60

Qy   61 CAGGGGCGTGGGGGGATTCCCCGTCTGCGAGAAGGGTGAGCCAACCACCTTAGTAT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CAGGGGCGTGGGGGGATTCCCCGTCTGCGAGAAGGGTGAGCCAACCACCTTAGTAT 120

Qy   121 GTAGGCGGCGGGGACTCATGACGCTCGCGTCGCGTGTGGTGGTGGCGCCTTAGGCAGC 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GTAGGCGGCGGGGACTCATGACGCTCGCGTCGCGTGTGGTGGTGGCGCCTTAGGCAGC 180

Qy   181 CCTGATGGGCGTTCATGGGTTCGGTGGGCACTGTAGGGAAGACCGGGGACCGG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCTGATGGGCGTTCATGGGTTCGGTGGGCACTGTAGGGAAGACCGGGGACCGG 240

Qy   241 CCTCCCAGATAGACGCGGCAGCGGCACTGTAGGGAAGACCGGGGACCGG 300
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCCCAGATAGACGCGGCAGCGGCACTGTAGGGAAGACCGGGGACCGG 300

Qy   301 CAGACCTCTTTTGAGTATCACGCCTCCGGAGTAGTTGGGCAAGCCCACCTATATGT 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAGACCTCTTTTGAGTATCACGCCTCCGGAGTAGTTGGGCAAGCCCACCTATATGT 360

Qy   361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TGGGATGGTTGGGGTTAGCCATCCATACCGTACTGCCTGATAGGGTCCTTGCGAGGGAT 420

Db   8341 AGAATTAACATCATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGGCAAGCCTTA 8400
Qy   8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTGCCGAGGGTCTGGG 8460
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8401 CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTCTGCCGAGGGTCTGGG 8460

Qy   8461 ATACAACCCCAGTGCGTGGGTGATTGGGTATCTAATACATCACTACCCATGTTGTGGGT 8520
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8461 ATACAACCCCAGTGCGTGGGTGATTGGGTATCTAATACATCACTACCCATGTTGTGGGT 8520

Qy   8521 TAGCCGTGTGTGTGGCTGTCATTCATGGAGCAGATGCTCTTGAGGACAAACTTCCCGA 8580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8521 TAGCCGTGTGTGTGGCTGTCATTCATGGAGCAGATGCTCTTGAGGACAAACTTCCCGA 8580

Qy   8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAGAGATCTGCCAG 8640
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8581 GACTGTGACCTTTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAGAAGATCTGCCAG 8640

Qy   8641 CATCATTGCTGCTGTGTGCACGGTATTGAGGCTTTCGGTGTGGTGCCGCGTCACCAACCGCTGA 8700
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8641 CATCATTGCTGCTGTGTGCACGGTATTGAGGCTTTCGGTGTGGTGCCGCGTCACCAACCGCTGA 8700

Qy   8701 GATCCTCAGAGTTCCCAATCACTACAGACATGACATGCCCCCCTGCCCTGGCG 8760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8701 GATCCTCAGAGTTCCCAATCACTACAGACATGACATGCCCCCCTGCCCTGGCG 8760

Qy   8761 AAAGAAAGCCAGGGCGGTCCTCGCCAGCGACCAAGAGCGTGGCGGAGCACACGCAAAATT 8820
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8761 AAAGAAAGCCAGGGCGGTCCTCGCCAGCGACCAAGAGCGTGGCGGAGCACACGCAAAATT 8820

Qy   8821 GGCTCGCTTCCTTCCTTCCTCTCGGCATGCTACATCTAGACCTTCTACCAGATTGGATAAGACGAG 8880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8821 GGCTCGCTTCCTTCCTTCCTCTCGGCATGCTACATCTAGACCTTCTACCAGATTGGATAAGACGAG 8880

Qy   8881 CGTGGCTTCGGTACACCACTTCAATTGTTACTCCCCGAGGGGGATGTGTT 8940
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db   8881 CGTGGCTTCGGTACACCACTTCAATTGTTACTCCCCGAGGGGGATGTGTT 8940

Qy   8941 TATTACCACCAGAGAGATTGCAGAGATTCCTTGTGAAGTATTTGGCTGTCATTGTTTT 9000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8941 TGTTACCACCAGAGAGAAGAGATTGCAGAGATTTCCTTGTGAAGTATTTGGCTGTCATTGTTTT 9000

Qy   9001 TGCCCTAGGGCTCATTGCTGTTGGATTAGCCATCAGCTGAACCCCCAAAATTCAAAATTAA 9060
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9001 TGCCCTAGGGCTCATTGCTGTTGGACTAGCCATCAGCTGAACCCCCAAAATTCAAAATTAA 9060

Qy   9061 CTAACAG----TTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC 9116
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   9061 TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAGCGGCAACAGGGGAGACCCC 9120

Qy   9117 GGGCTTAACGACCCCGC 9133
              |||||||||||||||||
Db   9121 GGGCTTAACGACCCCGC 9137

RESULT 4
US-08-424-550B-393
; Sequence 393, Application US/08424550B
; Publication No. US20020119447A1
; GENERAL INFORMATION:
;   APPLICANT: JOHN N. SIMONS
;   APPLICANT: TAMI J. PILOT-MATIAS
;   APPLICANT: GEORGE J. DAWSON
;   APPLICANT: GEORGE G. SCHLAUDER
;   APPLICANT: SURESH M. DESAI
;   APPLICANT: THOMAS P. LEARY
;   APPLICANT: ANTHONY SCOTT MUERHOFF
;   APPLICANT: JAMES C. ERKER
;   APPLICANT: SHERI L. BUIJK
;   APPLICANT: ISA K. MUSHAHWAR
;   TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
;   TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
;   NUMBER OF SEQUENCES: 716

```
Qy    421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACC  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGGGAGTCTCGTAGACCGTAGCACATGCCTGTTATTTCTACTCAAACAAGTCCTGTACC  480

Qy    481 TGCGCCCAGAGAACGCGCAAGAACAAGCAGAGCGCAGGCTTCATATCCTGTGTCATTAAAAC  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGCGCCCAGAGAACGCGCAAGAACAAGCAGAGCGCAGGCTTCATATCCTGTGTCATTAAAAC  540

Qy    541 ATCTGTTGAAAGGGACAACACGCAAAGCGCAAAGCGCAGCCGATGCTCGGCCTCGTAA  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ATCTGTTGAAAGGGACAACACGCAAAGCGCAAAGCGCAGCCGATGCTCGGCCTCGTAA  600

Qy    601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTACAAAATTGCTGGTATCCATGATGGCTTGCAGACATTGGCTCAGGCTGCTTTGCCAGC  660

Qy    661 TCATGGTTGGGACGGCGTCGTTGCACCAGTCTGCCAGATAGTACGGTTCTGGAATCTTCTGGA  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TCATGGTTGGGACGGCGTCGTTGCACCAGTCTGCCAGATAGTACGGTTCTGGAATCTTCTGGA  720

Qy    721 TTACCCTTTGGGTGGATTGGTGTGATGTTACAACTCACACACCTCTAGGACCCGCTGGT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TTACCCTTTGGGTGGATTGGTGTGATGTTACAACTCACACACCTCTAGGACCCGCTGGT  780

Qy    781 GGAGGAGGCGGTCGTTGCACCAGTCTGCCAGATAGTACGGTTCTGTCCACCTTTTGTGTCTGCTCGTCTTTGGCCTG  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GGAGGAGGCGGTCGTTGCACCAGTCTGCCAGATAGTACGGTTCTGTCCACCTTTTGTGTCTGCTCGTCTTTGGCCTG  840

Qy    841 CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTGTGTCTGCTATCTTTGGCCTG  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTGGGCTACTGGTTGGTTCGGTGTCCACCTTTTGTGTCTGCTATCTTTGGCCTG  900

Qy    901 TCCCTGTAGTAGTGGGGCGGCGGCCGGTCACTGACCGCAGACACAAATACCACCAATTG  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCCCTGTAGTAGTGGGGCGGCGGCCGGTCACTGACCGCAGACACAAATACCACCAATTG  960

Qy    961 CTGCCAGGCTAATCAGTTATCTATTGTTCTCCTTGCCACTGCCGAGCCTGGTTG  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CTGCCAGGCTAATCAGTTATCTATTGTTCTCCTTGCCACTGCCGAGCCTGGTTG  1020

Qy   1021 TGTGATCTGTGCCGGACGAGTGCTGGGTTCCCGCCAATCGTAGCATCCGTACACACCCTTCAA  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TGTGATCTGTGCCGGACGAGTGCTGGGTTCCCGCCAATCGTAGCATCCGTACACACCCTTCAA  1080

Qy   1081 TTGGGACTGGTTGGTCACCGACTTCCTTCTTGGCTGACCACCATTGTTGTTATGGGGCGTCTCTGT  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TTGGGACTGGTTGGTCACCGACTTCCTTCTTGGCTGACCACCATTGTTGTTATGGGGCGTCTCTGT  1140

Qy   1141 GACCTGTGACGCGCCCTTGACATTGGTGGTGTCGTGTATTAGTCGGTGACTG  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTGTGACGCGCCCTTGACATTGGTGGTGTCGTGTATTAGTCGGTGACTG  1200

Qy   1201 GCTTGTCAGGCACTGGTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GCTTGTCAGGCACTGGTTATTCACATAGACCTCAATGAAACTGGTACTTGTTACCTGGA  1260

Qy   1261 AGTTGCCCACTGGAATAGATCCTGGGTTTATCGGTGGATGCGGCCGGCAAGGT  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGTTGCCCACTGGAATAGATCCTGGGTTTATCGGTGGATGCGGCCGGCAAGGT  1320

Qy   1321 CGAGGCTGTCATCTTCTTGACCAAACTGCTTGCGCTCGATCTGCTCCATATCGCGACTAT  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CGAGGCTGTCATCTTCTTGACCAAACTGCTTGCGCTCGATCTGCTCCATATCGCGACTAT  1380

Qy   1381 GTTTAGCAGTGTACACTACCTGCGGCTCTGATCTGCTCTATGCCTCTCGGGGCAA  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GTTTAGCAGTGTACACTACCTGCGGCTCTGATCTGCTCTATGCCTCTCGGGGCAA  1440

Qy   1441 GTGGTATCAGTTGCTCTCCAGCGCTTATGCTTTACATAGAAGCGACCTCTGGAAACCCTAT  1500
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   1441 GTGGTATCAGTTGCTCTCCAGCGCTTATGCTTTACATAGAAGCGACCTCTGGAAACCCTAT  1500

Qy   1501 CAGGGTGCCCACTTGGATGCTCAATAGCTCAGTTTTGCTCGCCTTTGATGATACCATGTCC  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CAGGGTGCCCACTGGATGCTCAATAGCTCAATGAGTTTGCTCGCCTTTGATGATACCATGTCC  1560

Qy   1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC  1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TTGCCACTCTTATTTGAGTGAGAATGTGTCAGAAGTCATTGTTACAGTCCAAAGTGGAC  1620

Qy   1621 CAGGCCTATCACTCTAGATATAACAACTCCATATCTTGGTACCCCTATACAATCCCTG  1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CAGGCCTGTCACTCTAGATATAACAACTCCATATCTTGGTACCCCTATACAATCCCTG  1680

Qy   1681 TGGCAGGGGATGTGTGGTTAAATTCAAAAATAACACACATGGGGTTGCTGCCGTATTCGCAA  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGGCAGGGGATGTGTGGTTAAATTCAAAAATAACACACATGGGGTTGCTGCCGTATTCGCAA  1740

Qy   1741 TGTGCCATCGTACTGCACTATGGCAGTGTGGAACGCACACTCGCAACACTTA  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TGTGCCATCGTACTGCACTATGGCAGTGTGGAACGCACACTCGCAACACTTA  1800

Qy   1801 CGAAGCATGCGGTGTAACACCATGGCTACAACACCGCACAACGGCTCAGCCTGAA  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CGAAGCATGCGGTGTAACACCATGGCTACAACACCGCACAACGGCTCAGCCTGAA  1860

Qy   1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAAATGTTTAAACCTCATAATTGGATGTC  1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 ATTGGCTATATTACAATACCCTGGGTCTAAAGAAAATGTTTAAACCTCATAATTGGATGTC  1920

Qy   1921 AGGCCATTTGTATTTGAGGGATCAGATACCCCTAGTTTATTTTATGACCCTGTGAA  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AGGCCATTTGTATTTGAGGGATCAGATACCCCTAGTTTATTTTATGACCCTGTGAA  1980

Qy   1981 TTCCACTCTTCCTACCACCGGAGGAGTGGGCTAGGTTGCCCGGTACCCACCTGTGGTACG  2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TTCCACTCTTCCTACCACCGGAGGAGTGGGCTAGGTTGCCCGGTACCCACCTGTGGTACG  2040

Qy   2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG  2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TGGTTCTTGGTTACAGGTTCCGCAAGGGTTTTACAGTGATGTGAAAGACCTAGCCACAGG  2100

Qy   2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGTCTTATATTCCCACGGGTGC  2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ATTGATCACCAAAGACAAAGCCTGGAAAAATTATCAGTCTCTATATTCCCACGGGTGC  2160

Qy   2161 TTTGTCTTCACGGGAGTTACCACCAAGGCCGTGGTGCTAATTCTGTTGGGGTTGTGTGG  2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TTTGTCTTCACGGGAGTTACCACCAAGGCCGTGGTGCTAATTCTGTTGGGGTTGTGTGG  2220

Qy   2221 CAGCAAGTATCTTATTTAGCCTTATTTTAGCCTGTCCCTTGTTGTCCCCAAGCTGGCCGCTTC  2280
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   2221 CAGCAAGTATCTTATTTTAGCCTTATTTTAGCCTGTCCCTTGTTGTCCCCAAGCTGGCCGCTTC  2280

Qy   2281 TGGTTACCCTTTGGCTCCTGTGTCTCCCATCCGTATCGTCAGTCGTTGGGATGTG  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 TGGTTACCCTTTGGCTCCTGTGTCTCCCATCCGTATCGTCAGTCGTTGGGATGTG  2340

Qy   2341 TTTGTCTAAAGCTCAAGTAGCTCAAGTAGCTCCGTTTGCTTTCTTCATCTGTTGCTATCTCCG  2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   2341 TTTGTCTAAAGCTCAAGTAGCTCAAGTAGCTCCGTTTGCTTTCTTCATCTGTTGCTATCTCCG  2400

Qy   2401 CTGCAGGCTACGTTATGCGCGCATTTTGTGCCCATGCCATGCTGTGCCCCCT  2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 CTGCAGGCTACGTTATGCGCGCATTTTGTGCCCATGCCATGCTGTGCCCCCT  2460

Qy   2461 AACTTTCTTTGTTGCAGCAGCTGCGCCAACCAGATTATGACTGGTGCGGTGCGACTGCT  2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 AACTTTCTTTGTTGCAGCAGCTGCGCCAACCAGATTATGACTGGTGCGGTGCGACTGCT  2520

Qy   2521 AGTGGCAGGTAGTTTTGTGGGCCGGCGTAACCGTGGTCACCGGCATAGCTCTTGT  2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 AGTGGCAGGTAGTTTTGTGGGCCGGCGTAACCGTGGTCACCGGCATAGCTCTTGT  2580

Qy   2581 AGGTCCTTGGCCTCTTCGGTTAGCGCTTTAACCCTCTTGGTTGCATTGGTTACGGCGTCAGC  2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  2581 AGGTCCTTGGCCTCTCGGTAGCGCTTTAACCCTCTGCATTGGCTACGCCTGCTTCAGC 2640

Qy  2641 TTTTGATACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT 2700
         |||||  |||||||||||||||||||||||||||||||||||||||||||||||
Db  2641 TTTTGACACCGAGATAATTGGAGGGCTGACAATACCACCTGTAGTAGCATTAGTTGTCAT 2700

Qy  2701 GTCTCGTTTGGCTTCTTGCTCACTGTTACCTCGCTGGTGCTTTAGTTAACTCCTATCT 2760
         |||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 GTCTCGTTTGGCTTCTTGCTCACTGTTACCTCGCTGGTGCTTTAGTTAACTCCTATCT 2760

Qy  2761 TTGGCAACGTTGGGAGGAATTGGTTTTGGAACGTTACACTAACTAAGACGAGAGGTTTTCCT 2820
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 TTGGCAACGTTGGGAGGAATTGGTTTTGGAACGTTACACTAACTAAGACGAGAGGTTTCTCCT 2820

Qy  2821 TGTGTGTTGTTTCCCCGGTGCGACAATATGACGGCCGCTGGTGACTTCTGTGTGTGTCA 2880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821 TGTGTGTTGTTTCCCCGGTGCGACAATATGACGGCCGCTGGTGACTTCTGTGTGTGTCA 2880

Qy  2881 CGTAGCTCTCTATGTTAACATCCAGTGCAGCATGCCACATCCCAGTACCAGTCTCAGGGT 2940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881 CGTAGCTCTCTATGTTAACATCCAGTGCAGCATGCCACATCCCAGTACCAGTCTCAGGGT 2940

Qy  2941 TAGGGCCCATAGAATGTTGGTTGCGTTCGGAAAGTGTCATGTCTTGGTGTATTCTCATTATGT 3000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 TAGGGCCCATAGAATGTTGGTTGCGTTCGGAAAGTGTCATGTCTTGGTGTATTCTCATTATGT 3000

Qy  3001 TCTTAAGTTTTTCCCTCTAGTTGTTGGTGAGAATGGTGTGTTTTTCTATAAGCACTTGCA 3060
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 TCTTAAGTTTTTCCCTCTAGTTGTTGGTGAGAATGGTGTGTTTTTCTATAAGCACTTGCA 3060

Qy  3061 TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCATTTTTCCC 3120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 TGGTGATGTCTTGCCTAATGATTTTGCCTCGAAACTACCATTGCAAGAGCCATTTTTCCC 3120

Qy  3121 TTTTGAAGGCAAGGCCTCTATAGGAATGAAGAAGAGACGCTTGGCGTGTGGGGACAC 3180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 TTTTGAAGGCAAGGCCTCTATAGGAATGAAGAAGAGACGCTTGGCGTGTGGGGACAC 3180

Qy  3181 GGTTGATGGTTTGCCCGTTGTTGCGGCTCTCGGCGACCTTGTTTCGCAGGGTTGGCTAT 3240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 GGTTGATGGTTTGCCCGTTGTTGCGGCTCTCGGCGACCTTGTTTCGCAGGGTTGGCTAT 3240

Qy  3241 GCCGCCAGATGGGTGGGCCATTACCGGCCACCCTTTACGCTGCGAGTCTCTCTGAACGTGG 3300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 GCCGCCAGATGGGTGGGCCATTACCGGCCACCCTTTACGCTGCGAGTCTCTCTGAACGTGG 3300

Qy  3301 CACGCTGTCAGCGATGGCGATGGCCACTGGTCATGAGCGTATAGACCCCGAACTTGGACTGGAAC 3360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 CACGCTGTCAGCGATGGCGATGGCCACTGGTCATGAGCGTATAGACCCCGAACTTGGACTGGAAC 3360

Qy  3361 TATCTTCAGATTAGGATCTCTGGCCACTAGCTACATGGATTTGTTTGTGACAACGTGTT 3420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 TATCTTCAGATTAGGATCTCTGGCCACTAGCTACATGGATTTGTTTGTGACAACGTGTT 3420

Qy  3421 GTATACTGCTCACCATGGCAGCAAGGGCGCCGGTTGGCTCATCCCACCACAGGCTCTATACA 3480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421 GTATACTGCTCACCATGGCAGCAAGGGCGCCGGTTGGCTCATCCCACCACAGGCTCCATACA 3480

Qy  3481 CCCAATAACCGTTGACGCGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG 3540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481 CCCAATAACCGTTGACGCGCTAATGACCAGGACATCTATCAACCACCATGTGGAGCTGG 3540

Qy  3541 GTCCCTTACTCGTGTCTGTCTTGCCGTGGGAGACCAAGGGTATCTGGTAACACGACTGGGGTC 3600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541 GTCCCTTACTCGTGTCTGTCTTGCCGGGGGAGACCAAGGGTATCTGGTAACACGACTGGGGTC 3600

Qy  3601 ATTGGTTGAGGTCAACAAATCCGATGACCCCTTATTGGTGTGTGTGTGCGGGGGCCCTTCCCAT 3660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601 GTCCCTTACTCGTGTCTGTCTTGCCGATGACCCCTTATTGGTGTGTGTGTGCGGGGGCCCTTCCCAT 3660

Qy  3661 GGCTGTTGCCAAGGGTTCTTCAGGTTCTTCAGGTGCTGCCCCCGATTCTGTGCTTCTTGTGCTCATGTTATTGG 3720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661 GGCTGTTGCCAAGGGTTCTTCAGGTGCCCCTTCGGGTGCCCCGGATTCTGTGCTTCTTCCCCGGGCATGTTATTGG 3720
```

```
Qy  3721 GATGTTCACCGCGGTGCTAGAAATTCTGCGGTTCAGTCAGTCAGATTAGGTAGGCCGTT 3780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721 GATGTTCACCGCGGTGCTAGAAATTCTGCGGTTCAGTCAGTCAGATTAGGTAGGCCGTT 3780

Qy  3781 GGTGTGTGCTGATACCATCCCAGTACACAGCACATGCCACTCTTGTATACAAAACCTAC 3840
         ||||||||||||||||||||||||||||||||||||||||||||| ||||||  |||||
Db  3781 GGTGTGTGCTGATACCATCCCAGTACACAGCACATGCCACTCTTGATACAAAAACCTAC 3840

Qy  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGCGAAGTCAAC 3900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 TGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCGGCAAGTCAAC 3900

Qy  3901 CAAATTACCACTTCTTACATGCAGGAGAAGTATGAGTCTTGGTCCTAAATCCCAGTGT 3960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 CAAATTACCACTTCTTACATGCAGGAGAAGTATGAGTCTTGGTCCTAAATCCCAGTGT 3960

Qy  3961 GGCTACAACAGCATCAATGCCAAAGTACAACAACAGGGGGCTTCACTCAGCTACGGCGTGAATCAAATTG 4020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 GGCTACAACAGCATCAATGCCAAAGTACAACAACAGGGGGCTTCACTCAGCTACGGCGTGAATCAAATTG 4020

Qy  4021 CTATTTTAATGGCAAATGTACCAACACAGGGGCTTCACTCAGTGAATAATCATTGTGACGAATGCCATGCAT 4080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 CTATTTTAATGGCAAATGTACCAACACAGGGGCTTCACTCAGTGAATAATCATTGTGACGAATGCCATGCAT 4080

Qy  4081 GTACCTGACCGGACATGTTCCCGGAACTATGACGTCATCATTGTGACGCACTTGCA 4140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 GTACCTGACCGGACATGTTCCCGGAACTATGACGTCATCATTGTGACGCACTTGCA 4140

Qy  4141 TACCGATGCAACCACCGTGTTCTTGCCACGGTCCTAACCGAAGGTCCATCCAAAAA 4200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 TACCGATGCAACCACCGTGTTCTTGCCACGGTCCTAACCGAAGGTCCATCCAAAAA 4200

Qy  4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCTGGAGTAATCCTCACCACATGC 4260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201 TGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCCCTGGAGTAATCCTCACCACATGC 4260

Qy  4261 CAACATAACTGACTGAGATTCAATTAACCGATGAAGAACGATGAAAGAT 4320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261 CAACATAACTGACTGAGATTCAATTAACCGATGAAGAACGATGAAAGAT 4320

Qy  4321 TAAGGAGGAAAATCTGAAGAAGGGAGACACCTTATCTTTGAGGCTACCAAAAACACTG 4380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321 TAAGGAGGAAAATCTGAAGAAGGGAGACACCTTATCTTTGAGGCTACCAAAAAACACTG 4380

Qy  4381 TGATGAGCTTGCTAACGAGTAGCTGAAACAGCTGTCTCTTACTAGTTGCCACTGATGCCTTGTG 4440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381 TGATGAGCTTGCTAACGAGTAGCTGAAACAGCTGTCTCTTACTAGTTGCCACTGATGCCTTGTG 4440

Qy  4441 ATGTGACATCTCAAAAATCCCTGAGGGCGACTGTGTAGTTGCCACTGATGCCTTGTT 4500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441 ATGTGACATCTCAAAAATCCCTGAGGGCGACTGTGTAGTTGCCACTGATGCCTTGTT 4500

Qy  4501 TACAGGGTACACTGGTGACTTTGATTCCGTGATGGTCAGCGCTCATGGTAGAAGGCAC 4560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501 TACAGGGTACACTGGTGACTTTGATTCCGTGATGGTCAGCGCTCATGGTAGAAGGCAC 4560

Qy  4561 ATGGCCATGTTGACCTTGACCCTACTTTCACCATGGGGTGTTCGTGTGCGGGGTTTCAGC 4620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561 ATGGCCATGTTGACCTTGACCCTACTTTCACCATGGGGTGTTCGTGTGCGGGTCTCAGC 4620

Qy  4621 AATAGTTAAAGGCCAGCGGCCGCCGCACAGCCGTGGGCTGCATATACTACTA 4680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621 AATAGTTAAAGGCCAGCGGCCGCCGCACAGCCGTGGGCTGCATATACTACTA 4680

Qy  4681 TGTAGACGGGAGTTGTACCCCTTCGGGGTATGGTTCGTCATCAACAGAAGCTCAAGCCTTT 4740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681 TGTAGACGGGAGTTGTACCCCTTCGGGGTATGGTTCGTCATCAACAACAGAAGCTCAAACCTT 4740

Qy  4741 CGACGCAGCCAAGGCAATGGTATGGTTTGTCATCAACAGAAGCTCAAACTATTCTGGACAC 4800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741 CGACGCAGCCAAGGCAATGGTATGGTTTGTCATCAACAACAGAAGCTCAAACTATTCTGGACAC 4800
```

Right-hand alignment block:

```
Db  5881  TATGCTTGCTAGGAGCAACACTGTATGTAATGACTACTTTATTGCCACTCGTGACATCCG  5940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  5941  CAGGAAGATACTGGGCATTCTTGGAGGCATCTACCCCTGGAGTGTCATATCAGCTTGCAT  6000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5941  CAGGAAGATACTGGGCATTCTTGGAGGCATCTACCCCTGGAGTGTCATATCAGCTTGCAT  6000
Qy  6001  CCGTTGGCTCCACACCCGACGGAGGATGATTGCGGCCTCATTGCTTGGGGTCTCTAGAGAT  6060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6001  CCGTTGGCTCCACACCCGACGGAGGATGATTGCGGCCTCATTGCTTGGGGTCTCTAGAGAT  6060
Qy  6061  TTGGCAGTATGTGTGCAATTTCTTTGTGTGATTTGTGTGCAATTTCTTTAATGTCCTTAAAGCTGAGTTCA  6120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6061  TTGGCAGTATGTGTGCAATTTCTTTGTGTGATTTGTGTGCAATTTCTTTAATGTCCTTAAAGCTGAGTTCA  6120
Qy  6121  GAGCATGGTAACATTCCTGGTGTCCTTTCACAGCTGCCAGAAGGGGTACAACGGGCCC  6180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6121  GAGCATGGTAACATTCCTGGTGTCCTTTCACAGCTGCCAGAAGGGGTACAACGGGCCC  6180
Qy  6181  CTGATTGGATCAGGTATGCTCCAAGCACGCTGTCCATGCGGTGCTGGAACTCATCTTTTC  6240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6181  CTGATTGGATCAGGTATGCTCCAAGCACGCTGTCCATGCGGTGCTGGAACTCATCTTTTC  6240
Qy  6241  TGTTGAGAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6241  TGTTGAGAATGGTTTGCAAAACTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAG  6300
Qy  6301  AGGGGCTGTTCAACGCTAGGCTGTGTGGGTCGGCTAGACCGAACCCAACTGATTG  6360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6301  AGGGGCTGTTCAACGCTAGGCTGTGTGGGTCGGCTAGACCGAACCCAACTGATTG  6360
Qy  6361  GACTAGTCTTGTCGTCAATTATGCGTTAGGGACTACTGTAAATATGAAAAATGGGAGA  6420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6361  GACTAGTCTTGTCGTCAATTATGCGTTAGGGACTACTGTAAATATGAAAAATGGGAGA  6420
Qy  6421  TCACATTTTGTTACAGCAGTATCCTCCAAATGTCGTGTTTCACCCAGGTGCCCCCAAC  6480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6421  TCACATTTTGTTACAGCAGTATCCTCCAAATGTCGTGTTTCACCCAGGTGCCCCCAAC  6480
Qy  6481  CTTGAGAGCTGCAGTGGCCGTGGACGGCGGTTCAGTGTTATCTAGGTGAGCCCAA  6540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6481  CTTGAGAGCTGCAGTGGCCGTGGACGGCGGTTCAGTGTTATCTAGGTGAGCCCAA  6540
Qy  6541  AACTCCTTGGACGACATCTGCTTGCTGTGACGGTTACGGTCCTGACGGTAAGGGTAAAACTGTTAA  6600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6541  AACTCCTTGGACGACATCTGCTTGCTGTGACGGTTACGGTCCTGACGGTAAGGGTAAAACTGTTAA  6600
Qy  6601  GCTTCCCTTCCGCGGTTGACGGTGCACACCTGGTGTGGCCATGCACACTTAAATTTGCGTGA  6660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6601  GCTTCCCTTCCGCGGTTGACGGTGCACACCTGGTGTGGCCATGCACACTTAAATTTGCGTGA  6660
Qy  6661  TGCACTTGAGACAAATGACTGTAATTCCACAAACAACACTCCTAGTGATGAAGCCGCAGT  6720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6661  TGCACTTGAGACAAATGACTGTAATTCCATAAACAACACTCCTAGTGATGAAGCCGCAGT  6720
Qy  6721  GTCCGCTCTTGTTTTCAAACAGGAGTTGCGGCGGTACAAACCAATTGCTTGAGGCAATTTC  6780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6721  GTCCGCTCTTGTTTTCAAACAGGAGTTGCGGCGGTACAAACCAATTGCTTGAGGCAATTTC  6780
Qy  6781  AGCTGGCGGTTGACACCACCAAAACTGCCAGCCCTCCATCCGAAGAGGTAGTGGTAAGAAA  6840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6781  AGCTGGCGGTTGACACCACCAAAACTGCCAGCCCTCCATCCGAAGAGGTAGTGGTAAGAAA  6840
Qy  6841  GCGGCAGTTGCCGGGCAAGAACTGGTTCGCTTCGCCTCCCCCTTGCCTCCCCCTTACCTCCC  6900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6841  GCGGCAGTTGCCGGGCAAGAACTGGTTCGCTTCGCCTCCCCCTTGCCTCCCCCTTACCTCCC  6900
Qy  6901  AGGAGTGTCATGTCCTGAAAGCTGAAAGGTCAACGAAGTGACCCGTTGAGGAGTGGTGAAACCT  6960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6901  AGGAGTGTCATGTCCTGAAAGCTGAAAGGTCAACGAAGTGACCCGTTGAGGAGTGGTCTTCAAACCT  6960
Qy  6961  CCCTCCCTTCACCACCTGTTCTACAGTTGTAATGACTACTTTATTGCCACTCGTGGGAGCCGGGTGA  7020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Left-hand alignment block:

```
Qy  4801  CTATCGGCACCCAACCTGGGTTACCTGCGATAGAGCAACAAATTGGACGAGTGGCTGATCT  4860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801  CTATCGGCACCCAACCTGGGTTACCTGCGATAGAGCAACAAATTGGACGAGTGGCTGATCT  4860
Qy  4861  CTTTTCTATGGTCAACCCGACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA  4920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861  CTTTTCTATGGTCAACCCGACCTTCATTTGTCAATACTGCAAAAGAACTGCTGACAA  4920
Qy  4921  TTATGTTGTTGACTGCAGCCAACTACAACTGTCATCAGTGTATGCTATGCTGCTCC  4980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921  TTATGTTTGTTGACTGCAGCCAACTACAACTGTCATCAGTGTATGCTATGCTGCTCC  4980
Qy  4981  CAATGACGCACCACGGTGGCAGGGAGCCCGGCTTGGGAAAAACCTTGTGGGGTTCTGTG  5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981  CAATGACGCACCACGGTGGCAGGGAGCCCGGCTTGGGAAAAACCTTGTGGGGTTCTGTG  5040
Qy  5041  GCGCTTGGACGGCGCTGACGCCGTCTCCTGGCCCAGCCCCAGCGCGAGGGTGACCAGATACCA  5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041  GCGCTTGGACGGCGCTGACGCCGTCTCCTGGCCCAGCCCCAGCGCGAGGGTGACCAGATACCA  5100
Qy  5101  AATGTGCTTCACTGAAGTCTTATCTAGCCATTGACACTTTCTGGGCCCACTCGCTGTGGCGGTTGCTGTC  5160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5101  AATGTGCTTCACTGAAGTCTTATCTAGCCATTGACACTTTCTGGGCCCACTCGCTGTGGCGGTTGCTGTC  5160
Qy  5161  GGCTATGCTTATCTAGCCACCGGTGCTACTGTCTCCCCAGTGGTTGACGAAGAGAAATCGT  5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161  GGCTATGCTTATCTAGCCACCGGTGCTACTGTCTCCCCAGTGGTTGACGAAGAGAAATCGT  5220
Qy  5221  TATTACATCAGTCCCTCATGCAGCTACAAATCCTTGCTATCATAGAGTATTGCTTGGTTAGT  5280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221  TATTACATCAGTCCCTCATGCAGCTACAAATCCTTGCTATCATAGAGTATTGCTTGGTTAGT  5280
Qy  5281  GGAGGAGTGTGCATCATTCCCTTGCATGCGGCCATGTGACTTCGCAATTGACAAGCTGAA  5340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281  GGAGGAGTGTGCATCATTCCCTTGCATGCGGCCATGTGACTTCGCAATCGATAAGCTGAA  5340
Qy  5341  GAGTACAATCACCACAACTAGTCCTTTCACATTGGAAAACCGCCCCTTGAAAACTTAACAC  5400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341  GAGTACAATAACCACAACTAGTCCTTTCACATTGGAAAACCGCCCCTTGAAAACTTAACAC  5400
Qy  5401  CTTTCTTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGTTAGT  5460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401  CTTTCTTTGGGCCTCATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTGTTAGT  5460
Qy  5461  CACTTTACCTGACAATCGGTGGGTTTGTCTTTTGCTTCATTGCGTGTTTGCTTTCATTGGGTATTACTAC  5520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461  CACTTTACCTGACAATCGGTGGGTTTGTCTTTTGCTTCATTGCGTGTTTGCTTTCATTGGGTATTACTAC  5520
Qy  5521  CCCACTACCTCACAGAGATCAAAATGTTCCTGTCATTATTGGAGGCGCAATTGCGTCCAA  5580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521  CCCACTACCTCACAGAGATCAAAATGTTCCTGTCATTATTGGAGGCGCAATTGCGTCCAA  5580
Qy  5581  GCTTACAGACGTCTAGAGGCGCACTGAAGGCTCGGGCGTTCGCGGCGCTGCGCGGACCAGTCT  5640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581  GCTTACAGACGTCTAGAGGCGCACTGAAGGCTCGGGCGTTCGCGGCGCTGCGCGGACCAGTCT  5640
Qy  5701  TGGTACATGGACATCGGTGGGTTTGTCTTTTGCTTTGACATGGGTCGCCCACTATGGATCAGCT  5760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701  TGGTACATGGACATCGGTGGGTTTGTCTTTTGCTTTGACATGGGTCGCCCACTATGGATCAGCT  5760
Qy  5701  ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGAGGCCGCAGGAGTGCATCAGCT  5760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5701  ATCCACTGCTTGCTTGACATTTAAATGCTTGATGGGTGAGTGAGGCCGCAGGAGTGCATCAGCT  5760
Qy  5761  TGCTGGTTAGTCTACTTCCGCGGTTCAATCCGGTTCTGGGCGTTGGGCGCAGGAGTGCTGGGCGTCTTGTCAGC  5820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5761  TGCTGGTTAGTCTACTTCCGCGGTTCAATCCGGTTCTGGGCGTTGGGCGCAGGAGTGCTGGGCGTCTTGTCAGC  5820
Qy  5821  TTGTGCAATGTTTGCTTTGACAACAGCAGGGCCCAGATCACTGCCCAACAGACTTCTTAC  5880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5821  TTGTGCAATGTTTGCTTTGACAACAGCAGGGCCCAGATCACTGCCCAACAGACTTCTTAC  5880
Qy  5881  TATGCTTGCTAGGAGCAACACTGTATGTAATGACTACTTTATTGCCACTCGTGACATCCG  5940
```

```
Db  6961  CCCTTCTTCACCACTGTTCTACAGTTGGCCATGCCGATGCCCCTGTTGGGAGCAGTGA  7020

Qy  7021  GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAACAGGCGGAGGCCCTGATGA  7080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7021  GTGTAACCCTTTCACTGCAATTGGATGTGCAATGACCGAAACAGGCGGAGGCCCTGATGA  7080

Qy  7081  TTTACCCAGTTACCCTCCCAAAAAGGAGGTCTCTGAATGTCAGACGAAGTTGGTCGAC    7140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7081  TTTACCCAGTTACCCTCCCAAAAAGGAGGTCTCTGAATGGTCAGACGAAGTTGGTCAAC   7140

Qy  7141  GGCTACAACCGCTTCCAGCTTCCAGCTTCACTGGCCCCCGTACCCTAAGATACGGGAAAGA  7200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7141  GACTACAACCGCTTCCAGCTTCCAGCTTCACTGGCCCCCGTACCCTAAGATACGGGAAAGA  7200

Qy  7201  TTCCACTCAGTCAGCCCCGCCAAACGGCTACACCTGGACACGGCCCAAACGGCTACACCTG  7260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7201  TTCCACTCAGTCAGCCCCGCCAAACGGCTACACCTGGACACGGCCCAAACGGCTACACCTG  7260

Qy  7261  TTCGTGCAGCATGAGCTGAGCTACACTCGGGCCCATCACTAGTGGTTTCCTCAAACAAAGG  7320
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  7261  TTCGTGCAGCATGAGCTGAGCTACACTCGGGCCCATCACTAGTGGTTTCCTCAAACAAAGG  7320

Qy  7321  TCTGTCTGCAACTCGGGGCCATCACTAGTGGTTTCCTCAAACAAAGATCATTGGTGTATGT  7380
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  7321  TCTGTCTGCAACTCGGGGCCATCACTAGTGGTTTCCTCAAACAAAGATCATTGGTGTATGT  7380

Qy  7381  GACTGAGCCGCGGGATGCGGAGCTTAGAACAAAAGTCACTATTAATAGACAACCTCT     7440
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7381  GACTGAGCCGCGGGATGCGGAGCTTAGAACAAAAGTCACTATTAATAGACAACCTCT     7440

Qy  7441  GTTCCCCCCATCATCACCACAAGCAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAGTTGT  7500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7441  GTTCCCCCCATCATCACCACAAGCAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAGTTGT  7500

Qy  7501  CGGTGTCATGTGGGACTATGATGAAGTAGCAGCTCACACGGCCCTCAAGTCTGCTAAGTC  7560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7501  CGGTGTCATGTGGGACTATGATGAAGTAGCAGCTCACACGGCCCTCAAGTCTGCTAAGTC  7560

Qy  7561  CCACATCACTGCCCTTCGGGGACTGATGTTCGTTCTGGAGCAGCCCGCAAGGCTGTGTCT  7620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7561  CCACATCACTGCCCTTCGGGGACTGATGTTCGTTCTGGAGCAGCCCGCAAGGCTGTGTCT  7620

Qy  7621  GGACTTGCAGAAGCTGTGTCGAGGCAGGTGAGATACCGGTCATTATCGGCAAACTGTGAT  7680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7621  GGACTTGCAGAAGCTGTGTCGAGGCAGGTGAGATACCGGTCATTATCGGCAAACTGTGAT  7680

Qy  7681  AGTTCCAAAGGAGGAGTCTTCGTGAAGACCAACAAAGAAGAAACCCCAAG           7740
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7681  AGTTCCAAAGGAGGAGTCTTCGTGAAGACCAACAAAGAAGAAACCCCAAG           7740

Qy  7741  GCTTATCTCGTACCCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGT  7800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7741  GCTTATCTCGTACCCCCACCTTGAAATGAGATGTGTTGAGAAGATGTACTACGGTCAGGT  7800

Qy  7801  TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTACGGGTTTGTAGATCCACGTAC  7860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7801  TGCTCCTGACGTAGTTAAAGCTGTCATGGGAGATGCGTACGGGTTTGTGTCGACCACGTAC  7860

Qy  7861  CCGTGTCAAGGCTCTGTTGTGTCGATGGTGTCACCCGATGCAGTCGGAGCCACATGCGATAC  7920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7861  CCGTGTCAAGGCTCTGTTGTGTCGATGGTGTCACCCGATGCGGAGCCACATGCGATAC  7920

Qy  7921  AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTGGAGACAGACATCTACTC   7980
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  7921  AGTGTGTTTGACAGTACCATCACACCCGAGGATATCATGGTGGAGACAGACATCTACTC   7980

Qy  7981  AGCAGCTAAACTCAGTGACCAACACACCCGAGGATATCATGGTGGAGACAGACAGTTATA  8040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7981  AGCAGCTAAACTCAGTGACCAACACACCCGAGGATATCATGGTGGAGACAGAGTTATA   8040

Qy  8041  CGCTGAGGAGCCACGATGATCGCTTATGATGGCCGAGAGATCGGATATCCGAGAGATCGTGTAGGTC  8100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8041  CGCTGAGGAGCCACGATGATCGCTTATGATGGCCGAGAGATCGGATATCCGAGAGATCGTGTAGGTC  8100
```

```
Db  8101  TTCCGGGGTCTCATACTACCTCCAAGTCTTCCAACAGTTGACCTGCTGGCTGAAGTAAATGC  8160

Qy  8101  TTCCGGGGTCTCATACTACCTCCAAGTCTTCCAACAGTTGACCTGCTGGCTGAAGTAAATGC  8160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8101  TTCCGGGGTCTCATACTACCTCCAAGTCTTCCAACAGTTGACCTGCTGGCTGAAGTAAATGC  8160

Qy  8161  TGCAGCCGAACAGGCTGCAATGACGGTGCAATGACCGAAACCCTCGCTTCCTTATTGCGGCGATGATTGCAC  8220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8161  TGCAGCCGAACAGGCTGCAATGACGGTGCAATGACCGAAACCCTCGCTTCCTTATTGCGGCGATGATTGCAC  8220

Qy  8221  CGTAATTTGGAAGAGCGCCGGAGCAGATGCAGCAAACAAGCAATGGCTGTCTTTGCTAG  8280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8221  CGTAATTTGGAAGAGCGCCGGAGCAGATGCAGCAAACAAGCAATGGCTGTCTTTGCTAG  8280

Qy  8281  CTGGATGAAGGTGATGGGTGCACCACAAGAGTTGTGCCTCAACCAAATACAGTTTGGA   8340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8281  CTGGATGAAGGTGATGGGTGCACCACAAGAGTTGTGCCTCAACCAAATACAGTTTGGA   8340

Qy  8341  AGAATTAACATCATGCTCATCAAATGTACCTCGGAATTACCAAAAGTGGCAAGCCTTA   8400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8341  AGAATTAACATCATGCTCATCAAATGTACCTCGGAATTACCAAAAGTGGCAAGCCTTA   8400

Qy  8401  CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTGCCTCGCCGAGGGTCTGGG  8460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8401  CTACTTTCTTACAAGAGATCCTCGTATCCCCCTTGGCAGGTGCTGCCTCGCCGAGGGTCTCGGG  8460

Qy  8461  ATACAACCCAGTGCTGCGTGACGATGTCTTTCATGGAGCAGATGATGCTCAATACACTTCCCGA  8520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8461  ATACAACCCAGTGCTGCGTGACGATGTCTTTCATGGAGCAGATGATGCTCAATACACTTCCCGA  8520

Qy  8521  TAGCCGTGTGTGGCTGTCGTGCCATTTCATGGAGCAAACTTCAAAAGTTGT           8580
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8521  TAGCCGTGTGTGGCTGTCGTGCCATTTCATGGAGCAAACTTCAAAAGTTGT           8580

Qy  8581  GACTGTGACCTTTGACTGGTATGGAGAAAAATATACGGTGCCTGTAGAAGATCTGCCCAG  8640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8581  GACTGTGACCTTTGACTGGTATGGAGAAAAATATACGGTGCCTGTAGAGATCTGCCCAG  8640

Qy  8641  CATCATTGCTGGTGTGCACGGTTCTTCCGGTGGTGCGCTACACCAACGCTGA         8700
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8641  CATCATTGCTGGTGTGCACGGTTCTTCCGGTGGTGCGCTACACCAACGCTGA         8700

Qy  8701  GATCCTCAGAGTTCCCAATCAACACAGACATGACCAACTGTGATCGGCAAACTGTGAT   8760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8701  GATCCTCAGAGTTCCCAATCAACACAGACATGACCAACTGTGATCGGCAAACTGTGAT   8760

Qy  8761  AAAGAAAGCCAGGCGGTCCTCTGCCCCTGCGGGAGCACACGGCACACGCCAAATT     8820
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8761  AAAGAAAGCCAGGCGGTCCTCTGCCCCTGCGGGAGCACACGGCACACGCCAAATT     8820

Qy  8821  GGCTCGCTTCCTCTCTGCCATGCTCACATCTAGACCTCTACCAGATTGGAATAAGACGAG  8880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8821  GGCTCGCTTCCTCTCTGCCATGCTCACATCTAGACCTCTACCAGATTGGAATAAGACGAG  8880

Qy  8881  CGTGGCTCGGTTACGCACACTTTCAATTATTGTGATGTGTTTACTCCCCGGGGGATGTGTT  8940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8881  CGTGGCTCGGTTACGCACACTTTCAATTATTGTGATGTGTTTACTCCCCGGGGGATGTGTT  8940

Qy  8941  TATTACACCACAGAGAAGAGATTGCAGAAGTTCCTTGTGTGAAGTATTTGGCTGTGTTTT   9000
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  8941  TGTTACACCACAGAGAAGAGATTGCAGAAGTTTCTTGTGTGAAGTATTTGGCTGTGTTTT   9000

Qy  9001  TGCCCTAGGGCTCATTGCTGTGTTGGATTAGCCATCAGTGAACCCCAAAATTCAAAAATTAA  9060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  9001  TGCCCTAGGGCTCATTGCTGTGTTGGACTAGCCATCAGTGAACCCCAAAATTCAAAATTAA  9060

Qy  9061  CTAACAG----TTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGCAACAGGGGAGACCCC  9116
          ||||||||   |||||||||||||||||||||||||||||||||||||||||||
Db  9061  TTAACAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGGCAGCGGCAACAGGGGAGACCCC  9120

Qy  9117  GGGCTTAACGACCCGC  9133
          ||||||||||||||||
Db  9121  GGGCTTAACGACCCGC  9137
```

```
RESULT 5
US-08-424-550B-11
; Sequence 11, Application US/08424550B
; Publication No. US20020119447A1
  GENERAL INFORMATION:
;  APPLICANT: JOHN N. SIMONS
;  APPLICANT: TAMI J. PILOT-MATIAS
;  APPLICANT: GEORGE J. DAWSON
;  APPLICANT: GEORGE G. SCHLAUDER
;  APPLICANT: SURESH M. DESAI
;  APPLICANT: THOMAS P. LEARY
;  APPLICANT: ANTHONY SCOTT MUERHOFF
;  APPLICANT: JAMES C. ERKER
;  APPLICANT: SHERI L. BUIJK
;  APPLICANT: ISA K. MUSHAHWAR
;  TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
;  TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
;  NUMBER OF SEQUENCES: 716
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
;  STREET: 100 ABBOTT PARK ROAD
;  CITY: ABBOTT PARK
;  STATE: IL
;  COUNTRY: USA
;  ZIP: 60064-3500
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: Floppy disk
;  COMPUTER: IBM PC compatible
;  OPERATING SYSTEM: PC-DOS/MS-DOS
;  SOFTWARE: PatentIn Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/08/424,550B
;  FILING DATE:
;  CLASSIFICATION: 435435
;  ATTORNEY/AGENT INFORMATION:
;  NAME: POREMBSKI, PRISCILLA E.
;  REGISTRATION NUMBER: 33,207
;  REFERENCE/DOCKET NUMBER: 5527.PC.01
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE: 708-937-6365
;  TELEFAX: 708-938-2623
  INFORMATION FOR SEQ ID NO: 11:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH: 8912 base pairs
;  TYPE: nucleic acid
;  STRANDEDNESS: single
;  TOPOLOGY: linear
;  MOLECULE TYPE: DNA (genomic)
US-08-424-550B-11

Query Match     92.5%;   Score 8692.6;   DB 8;   Length 8912;
Best Local Similarity  98.7%;   Pred. No. 0;
Matches 8781;   Conservative  56;  Mismatches  52;   Indels  10;  Gaps   8;
```

```
Qy    196 TGGGTTCGGTGGTGGTGCGCGCTTTAGGCAGCCTCCACGCCCACCTCCCAGATAGAGC 255
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11 TGGGTTCGGTGGTGGTGCGCGCTTTAGGCAGCCTCCACGCCCACCTCCCAGATAGAGC  70

Qy    256 GGCGGCACTGTAGGGAAGGAAGCACCGGGGGACACCGGGGGACACCGGGGGACACCGGGGGGACACCG 315
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     71 GGCGGCACTGTAGGGAAGGAAGCACCGGGGGACACCGGGGGACACCGGGGGACACCGGGGGGACACCG 130

Qy    316 GTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACCACCTTGGGATGGTTGGGGT 375
          |||||||||||||||||||||||||||||||||||||||||||
Db    131 GTATCACGCCTCCGGAAGTAGTTGGGCAAGCCCACCACCTAYATGTTGGGATGGTTGGGGT 190

Qy    376 TAGCCATCCATACCGTACCGTATCGCCTGATAGGGTCCTTGCCGAGGGGATCTGGGAGTCTCGTAG 435
          |||||||||||||||||||||||||||||||||
Db    191 TAGCCATCCATACCGTACCGTATCGCCTGATAGGGTCCTTGCCGAGGGGATCTGGGAGTCTCGTAG 250

Qy    436 ACCGTAGCACATGCCTGTTATTCTACTCAAACAAGTCCTGTACCTGCGCCCAGAACGCG
          ||||||||||||||||||||||||||||||||||||||||:||||||||||
```

Nucleotide sequence alignment (Qy vs Db) with coordinate positions.

Left column:

```
Db  3547 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         CTAGAAATTCTGGCGGTTCAGTCGGCCAGATTAGGGTTAGGCCGTTGGTGTGTCGTGGAT 3606

Qy  3795 ACCATCCCAGTACACAGCACATGCCACTCGCCACTCTTGATACAAAACTACTGTGCCTAACGAGT 3854
         |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  3607 ACCATCCCAGTACACAGCACATGCCACGCCACTCTTGATACAAAACCTACTGTGCCTAACGAGT 3666

Qy  3855 ATTCAGTGCAAATTTTAATTGCCCCCACTGGCAGCGGCAAGTCAACCAAATTACCACTTT 3914
         |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db  3667 ATTCAGTGCAAATTTTAATTGCCCCCACTGGCAGCGGCAAGTCAACCAAATTACCACTTT 3726

Qy  3915 CTTACATGCAGGAGGAGTATGAGGTCTTGGTCCTCCAGTGTGGCTACAACAGCAT 3974
         ||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db  3727 CTTACATGCAGGRGAAGYATGAGGTCTTGGTCCTCCAAATCCCGAGTGTGGCTACAACAGCAT 3786

Qy  3975 CAATGCCAAAGTACATGCACGGCGACGCGGCTACGGCGTGAATCCAAATTGCTATTTTAATGCA 4034
         ||||||||||||||| |||||| ::|||||||||||||||||||||||||||||||||||
Db  3787 CAATGCCAAAGTACATGCACGGCGACGCGGCTACGGCGTGAATCCAAATTGCTATTTTAATGCA 3846

Qy  4035 AATGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGGCATGTACCTGACCCGAG 4094
         ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db  3847 AATGTACCAACACAGGGGCTTCACTTACGTACAGCACATATGGCATGTACCTGACCCGAC 3906

Qy  4095 CATGTTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGCTACCGCCAACCA 4154
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3907 GATGTTCCCGGAACTATGATGTAATCATTTGTGACGAATGCCATGCTACCGCCGAACCA 3966

Qy  4155 CCGTGTGGGAAGGTCCTAACCAGAGCTCCATCCAAAAATGTTAGGCTAGTGG 4214
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3967 CCGTGTGGGAAGGTCCTAACCAGAGCTCCATCCAAAAATGTTAGGCTAGTGG 4026

Qy  4215 TTCTTGCCACGGCTACCCCCTGGAGTAATCCCACCACATGCCAACATAACTGAGA 4274
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4027 TTCTTGCCACGGCTACCCCCTGGAGTAATCCCACCACATGCCAACATAACTGAGA 4086

Qy  4275 TTCAATTAAACCGATGAACGGCACTATCCCTTTCATGTGAAAAAAGATTAAGGAGGAAAATC 4334
         ||||||||||||| ::|||||||||||||||||||||||||||||||||||||||||||||
Db  4087 TTCAATTAACYGATGAACGGCACTATCCCTTTCATGTGAAAAAAGATTAAGGAGGAAAATC 4146

Qy  4335 TGAAGAAAGGGAGACACACCTTATCTTCTTGAGGGCTACCAAGAACACTGTGATGAGCTTGCTA 4394
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4147 TGAAGAAAGGGAGACACACCTTATCTTCGAGCGCTACCAGCGCTGTGATGAGCTTGCTA 4206

Qy  4395 ACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTTCTTACTATAGGGGATGTGACATCTCAA 4454
         ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db  4207 ACGAGTTAGCTCGAAAGGGAATAACAGCTGTCTCTTTCTTACTATAGGGGATGTGACATCTCAA 4266

Qy  4455 AAATCCCTGAGGGCGACTGTGTAGTAGTTGCCACTGCCTTGTCGTGTGTACAGGGTACACTG 4514
         ||||||||||||||| ::|||||||||||||||||||||||||||||||||||||||||||
Db  4267 AAATGCCTGAGGGCGACTGTGTAGTAGTTGCCACTGCCTTGTCGTGTGTACAGGGTACACTG 4326

Qy  4515 GTACCCCTTCGGGTATGCGTGTATGACTGCAGCCTCATGGTAGAAGGCACATGCCATGTTGACC 4574
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4327 GTACCCCTTCGGGTATGCGTGTATGACTGCAGCCTCATGGTAGAAGGCACATGCCATGTTGACC 4386

Qy  4575 TTGACCCTACTTTCACCATGGGTGTCGTCGTGTGCGGGGTTTCAGCAATAGTTAAAGGCC 4634
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4387 TTGACCCTACTTTCACCATGGGTGTTCGTGTGCGGGGTTTCAGCAATAGTTAAAGGCC 4446

Qy  4635 AGCGTAGGGGCCCACAGGCCGCGGGAGAGCTGGCATATACTACTATGTAGACGGGGAGTT 4694
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4447 AGCGTAGGGGCCCACAGGCCGCGGGAGAGCTGGCATATACTACTATGTAGACGGGGAGTT 4506

Qy  4695 GTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGGCCTTCGCACGCAGCCAAGG 4754
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4507 GTACCCCTTCGGGTATGGTTCCTGAATGCAACATTGTTGAAGGCCTTCGCACGCAGCCAAGG 4566

Qy  4755 CATGGTATGGTTTGTCATCAACAGAAGCTCAAACTATTCTGGACACCTATCGCCACCCAAC 4814
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4567 CATGGTATGGTTTGTCATCAACAGAAGCTCAAACTATTCTGGACACCTATCGCCACCCAAC 4626

Qy  4815 CTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCTCTTTTCTATGGTCA 4874
         ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4627 CTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCTCTTTTCTATGGTCA 4686
```

Right column:

```
Db  4627 CTGGGTTACCTGCGATAGGAGCAAATTTGGACGAGTGGGCTGATCTCTTTTCTATGGTCA 4686

Qy  4875 ACCCGAACCTTCATTTGTCAATACTGCAAAAAGAAACTGCTGACAATTATGTTTGTTGA 4934
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4687 ACCCGAACCTTCATTTGTCAATACTGCAAAAAGAAACTGCTGACAATTATGTTTGTTGA 4746

Qy  4935 CTGCAGCCCAACTACAACTGTGTCATCAGTAGTGCTATGCTATGCTGTCGTCTGTCCCAATGACCGCACCAC 4994
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4747 CTGCAGCCCAACTACAACTGTGTCATCAGTAGTGCTATGCTATGCTGTCGTCTGTCCCCAATGACCGCACCAC 4806

Qy  4995 GGTGGCAGGGAGCCGGCTTGTGGGAAAAACCTTGTGGGGTTCTGTGGCCTTGGACGGGCG 5054
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4807 GGTGGCAGGGAGCCGGCTTGGGAAAAACCTTGTGGGGCCGTTGGGCGCTTGGACGGGCT 4866

Qy  5055 CTGACGCCTGTCCTGGCCTGGCCCCAGAGCCCAGCGAGTGACCAGATACCAAATGTGCTTCACTG 5114
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4867 GTGACGCCTGTCCTGGCCTGGCCCCAGAGCCCAGCGAGTGACCAGATACCAAATGTGCTTCACTG 4926

Qy  5115 AAGTCAATACTTCTGGACAGCGCCACTCGCTCGTCGTGTTGGGCGTTGGGGCGTTGGGCGGTATGGCTTATC 5174
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4927 AAGTCAATACTTCTGGACAGCGCCACTCGCTCGTCGTGTTGGGCGTTGGGGCGTTGGGCGGTATGGCTTATC 4986

Qy  5175 TAGCCATTGACACTTTTGGCGCCACTTTGGCGCCGGTTGCTGTCGTCTGTCTATTACATCAGTCC 5234
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4987 TAGCCATTGACACTTTTGGCGCCACTTTGGCGCCGGTTGCTGTCGTCTGTCTATTACATCAGTCC 5046

Qy  5235 CTACCGGTGCTACTGTCGCCCCCAGTGGTTGACGGTTGCTGCAATTGACAAGTGAAGAGTACAATCACCA 5294
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5047 CTACCGGTGCTACTGTCGCCCCCAGTGGTTGACGGTTGCTGCAATTGACAAGTGAAGAGTGTGCAT 5106

Qy  5295 CATTCATTCCCTTGGAGGCCATGCCATGCCTGAAAAACCGCCCTTGAAAAACTAACACCTTTCTTGGGCCTC 5354
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5107 CATTCATTCCCTTGGAGGCCATGCCATGCCTGAAAAACCGCCCTTGAAAAACTTAACACCTTTCTTGGGCCTC 5166

Qy  5355 CAACTAGTCCTTTCACATTGGAAACCGGCGTCGCGGGAACAGCTCTTGGTAGTCACTTACCTGACA 5414
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5167 CAACTAGTCCTTTCACATTGGAAACCGGCGTCGGYGGGAACAGCTCTTGGTAGTCACTTACCTGACA 5226

Qy  5415 ATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTTTCATTGCGGGTATTACTACCCCACA 5474
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5227 ATGCAGCTACAATCCTTGCTATCATAGAGTATTGCTGTTTCATTGCGGGTATTACTACCCCACA 5286

Qy  5475 ATCCCTTTGCATCATGCCGTGTTTGTTCATTGCGGGTATTACTACCCCACA 5534
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  5287 ATCCCTTTGCATCATGCCGTGTTTGTTCATTGCGGGTATTACTACCCCACA 5346

Qy  5535 AGATCAAAATGTTCCTGTCATTATTGGAGGCGCCAATTGCGTCCAAGCTTACAGACGCTA 5594
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5347 AGATCAAAATGTTCCTGTCATTATTGGAGGCGCCAATTGCGTCCAAGCTTACAGACGCTA 5406

Qy  5595 GAGGGCACTGGCCGTCATGATGGCCGGGGCTCGCGGAACAGCTCCTTGGTACATGGACAT 5654
         |||||||||||||||||||||||||||||||||||||| :|||||||||||||||||||
Db  5407 GAGRGCACTGGCCGTCATGATGGCCGYGGGCTCGYGGAACAGCTCTTGGTACATGGACAT 5466

Qy  5655 CGGTGGGTTTGTCTTTGTCTTGTTCTTCATTGCGCTATGCTGCCGCCTCATCCGCCTTGCT 5714
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5467 CGGTGGGTTTGTCTTTGTCTTGTTCTTCATTGCGCTATGCTGCCGCCTCATCCGCCTTGCT 5526

Qy  5715 TGACATTTAAATGCTTGAGGGTGATGGTGAGTGCGCCACTATGGTTTAGTCT 5774
         ||||||| |||||||||||||||||| :||||||||||||||||||||||||
Db  5527 TGACATTTAAATGCTTGAGGGTGATGGTGAGTGCGCYCACTATGGTTTAGTCT 5586

Qy  5775 ACTCCGCGTTCAATCCGGCTCTTGTCAGCGGCTTGTGTCGCTTGTCAATGTTTG 5834
         ||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  5587 ACTCCGCGTTCAATCCGGCTCTTGYGVGGCTTGTGTCGCTTGTCAATGTTTG 5646

Qy  5835 CTTTGACAACAGCAGGGGCCAGATCACTGCCCAACAGACTTCTTACTATGCTTGCTAGGA 5894
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5647 CTTTGACAACAGCAGGGGCCAGATCACTGCCCAACAGACTTCTTACTATGCTTGCTAGGA 5706

Qy  5895 GCAACACTGTATGTAATGAGTACTTTATTGCCCACTCGTGACATCCGCAGGAAGATACTGG 5954
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5707 GCAACACTGTATGTARTGAGTACTTTATTGCCCACTCGTGACATCCGCAGGAAGATACTGG 5766
```

Qy   5955 GCATTCTGGAGGCATTCACCCCCTGGAGTGTCATATCAGCTTGCATCGTTGGCTCCACA 6014
           ||||||||||||||||||||||| || ||||| |||||||||||||| |||||||||
Db   5767 GCATTCTGGAGGCATCACCCCCTCCACCCCTGGAGTRITCATATCAGCTTGCATCCGTTGGCTYCACA 5826

Qy   6015 CCCCGACGGAGGATGATTGCGGCCTCATTGCTTGGGGCTTAGAGATTGGCAGTATGTGT 6074
Db   5827 CCCCGACGGAGGATGATTGCGGCCTCATTGCTTGGGGCTCTARAGATTGGCAGTATGTGT 5886

Qy   6075 GCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGCTGGAGTTCAGAGAGCATGGTTAACA 6134
Db   5887 GCAATTTCTTTGTGATTTGCTTTAATGTCCTTAAAGCTGGAGTTCAGAGAGCATGGTTAACA 5946

Qy   6135 TTCCTGGTGTGTCCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCCCTGGATTGGATCAG 6194
Db   5947 TTCCTGGTTGTCCTTTCTACAGCTGCCAGAAGGGGTACAAGGGCCCCTGGATTGGATCAG 6006

Qy   6195 GTATGTCCAAGCACGCTGTCCATGCGGTCGGCTAGACCGGACCCAACTCATCTTTCTGTTGAGAATGGTT 6254
Db   6007 GTATGTCCAAGCACGCTGTCCATGCGGTGCGGCTAGACCGGACCCAACTCATCTTTCTGTTGAGAATGGTT 6066

Qy   6255 TTGCAAAACTTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAGCAACCCAGACCTGTTCCAG 6314
Db   6067 TTGCAAAACTTTACAAAGGACCCAGAACTTGTTCAAATTACTGGAGCAACCCTGTTCCAG 6126

Qy   6315 TCAACGTTAGGCGTGTGTGGGTCGGCTAGACCGGTGCGGCTAGACCGGTGCGGCTAGACCGG 6374
Db   6127 TCAACGTTAGGCGTGTGTGGGTCGGCTAGACCGGTGCGGCTAGACCGGTGCGGCTAGACCGG 6186

Qy   6375 TCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAATGGGAGATCACATTTTTGTTA 6434
Db   6187 TCAATTATGGCGTTAGGGACTACTGTAAATATGAGAAATGGGAGATCACATTTTTGTTA 6246

Qy   6435 CAGCAGTATCCTCCCAAATGTCTGTTCACCCAGTGCCCCAACCTTGAGAGCTGCAG 6494
Db   6247 CAGCAGTATCCTCCCAAATGTCTGTTCACCCAGTGCCCCAACCTTGAGAGCTGCAG 6306

Qy   6495 TGGCCGTGGACGGGCTACAGTTCAGTGTTATCTGTAGGTGAGCCCAAAACTCCTTGGACGA 6554
Db   6307 TGGCCGTGGACCGGCTACAGTTCAGYGTATCTGTAGGTGAGCCCAAAACTCCTTGGACGA 6366

Qy   6555 CATCTGCTTGCTTGTGTTACGGTCCTGACGTGTTAAGGGTAAAACTGTTAAGCTTCCCTTCCGCG 6614
Db   6367 CATCTGCTTGCTTGTGTTACGGTCCTGACGTGTTAAGGGTAAAACTGTTAAGCTTCCCTTCCGCG 6426

Qy   6615 TTGACGGTCACACACCTGGTGTGCGCATGCAACTTAATTTGCGTGATGCACTTGAGACAA 6674
Db   6427 TTGACGGACACACACCTGGTGTGCGCATGCAACTTAATTTGCGTGATGCGACTTGAGGCAA 6486

Qy   6675 ATGACTGTAATTCCACAAACACACTTCCTTACCTTGCCTCCCCCTCAGTGTCCGCTCTGTTT 6734
Db   6487 ATGACTGTAATTCCATAAACACACTTCCTTACCTTGCCTCCCCCTCAGTGTCCGCTCTGTTT 6546

Qy   6735 TCAAACAGGAGTTGCAACGAGTTGACGCGTTAGAAGGCAATTTCGTTAGGAGGTCCTTGACA 6794
Db   6547 TCAAACAGGAGTTGCGGCGTACAACGAGTGACGCGTTAGAAGGTGACTTGAGGCAA 6606

Qy   6795 CCACCCAACTGCCAGCCCCCTCC---ATCGAAGAGGTAGTGTAAGAAAGCGCCAGTTCC 6851
Db   6607 CCACCCAACTGCCAGCCCCCCCCGAGATCGAGATGGTGTAAGAAAGCGCCAGTTCC 6666

Qy   6852 GGGCAAGAAGTTGGTTCGTTACCTTGCCTCCCCCTCCGAGATCCGTCCCAGAGTGTCAT 6911
Db   6667 GGGCAAGAAGTTGGTTCGTTCCTTACCTTGCCTCCCCCTCCGAGATCCGTCCCAGAGTGTCAT 6726

Qy   6912 GTCCTGAAAGCTGCAACGAGTTGACGCGTTAGAAGGTCCTTCAAACCTTCCCTTCCTTCAC 6971
Db   6727 GTCCTGAAAGCTGCAACGAGTGACGCGTTAGAAGGTGACTTCAAMCCTTCCCTTCTTCAC 6786

Qy   6972 CACCTTGTTCTACAGTTGGCCATGCCGATGCCGATGCCGATGCCGTTGGGAGCGGGTGAGTGTAACCCTT 7031
Db   6787 CACCTTGTTCTRCAGTTGGCCATGCCGATGCCGATGCCGATGCCGTTGGGAGCGGCAGCAGTGAGTGTAACCCTT 6846

Qy   7032 TCACTGCAATTGGATGTGCAATGACCGAAACAGGCGGAGGCCCTGATGATTACCAGTT 7091
           ||||||||||||||||||||||||||||: :||||||| | |:||||||||||||
Db   6847 TCACTGCAATTGGATGTGCAATGACCGAAACARGYGGAGKCCCWMAKRATTTACCAGTT 6906

Qy   7092 ACCCTCCCAAAAGGAGGAGTCTCTGAATGGTCAGACGAAAGTTGGTCGACGAAAGTTGGTCAACGGCTACAACCG 7151
Db   6907 ACCCTCCCAAAAGGAGGAGTCTCTGAATGGTCAGACGAAAGTTGGTCAACGACTACAACCG 6966

Qy   7152 CTTCCAGTACGTTACTGCGCCCCGTACCTTAAGATACGGGAAAGATTCCACTCAGT 7211
Db   6967 CTTCCAGTACGTTACTGCGCCCCGTACCCTAAGATACGGGGCAAGGATTCCACTCAAT 7026

Qy   7212 CAGCCCCGCCAAACGGCCTACAAAAGGCCCTACAAAAAGAGAGTGGGAAAGAGTGAGTTTCGTGCAGCA 7271
Db   7027 CAGCCACCGCCCAAACGGCCTACAAAAGGCCCTACAAAGAGAGTGGGAAAGAGTGAGTTTCGTGCAGCA 7086

Qy   7272 TGAGCTACACCTGGACCGACGTGATTAGCTTCAAAACTGCTTCTAAAGTTCTGTGCAA 7331
Db   7087 TGAGCTACACTTGGACCGACGTGATTAGCTTCAAAACAGAGATCATTGGTGTATGGACTGAGCCGC 7146

Qy   7332 CTCGGGGCCATCACTAGTGTGTTCCTCAAACAAAAGTCACTATTAATAGACAACCTCTGTTCCCCCCAT 7391
Db   7147 CTCGGGGCCATCACTAGTGGTTTCCTCAAACAAAAGTCACTATTAATAGACAACCTCTGTTCCCCCCAT 7206

Qy   7392 GGGATGCGGAGCTTAGAAACAAAGATGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGTATGGACTTGTGGTCATGT 7451
Db   7207 GGGATGCGGAGCTTAGAAACAAAGATGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGTATGGACTGTCATGT 7266

Qy   7452 CATACCACAAGCAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGTCGGGTCATGT 7511
Db   7267 CATACCACAAGCAAGCAAGTGAGATTGGCTAAGGAAAAAGCTTCAAAAGTTGTCGGGTCATGT 7326

Qy   7512 GGGACTATGATGAAGTAGCAGTCACACGCCCTCTAAGTCTCGTCTAAGTCCTGCTGCAG 7571
Db   7327 GGGACTATGATGAAGTAGCAGTCACACGCCCTCTAAGTCTCGTCTAAGTCCCACATCACTG 7386

Qy   7572 GCCTTCGGGGCACTGATGTTCGTTCTGGAGCAGCCCGCAAGCGTGTTCTGGACTTGCAGA 7631
Db   7387 GCCTTCGGGGCACTGATGTTCGTTCTGGAGCAGCCCGCAAGCGTGTTCTGGACTTGCAGA 7446

Qy   7632 AGTGTGTCGAGGCAGGTGAGATACCGGCAAACTGCATTATCGGCAAACTGTGATAGTTCCAAAGG 7691
Db   7447 AGTGTGTCGAGGCAGGTGAGATACCGGCAAACTGCATTATCGGCAAACTGTGATAGTTCCAAAGG 7506

Qy   7692 AGGAGGTCTTCGTGAAGACCCCCCAGAGAAACCAACAAAGAAACCCCCCAAGGCTTATCTCGT 7751
Db   7507 AGGAGGTCTTCGTGAAGACCCCCCAGAGAAACCAACAAAGAAACCCCCCAAGGCTTATCTCGT 7566

Qy   7752 ACCCCACCTTGTTGTTGCTTCGTGCGGTGTTGCTCCTGACG 7811
Db   7567 ACCCCACCTTGTTGTTGCTTCGTGCGGTGTTGCTCCTGACG 7626

Qy   7812 ACAGTACCATCACACCCGAGATATCATGGTGGAGACAGACATCTACTTCAGCAGCTAAAC 7871
Db   7627 ACAGTACCATCACACCCGAGATATCATGGTGGAGACAGACATCTACTTCAGCAGCTAAAC 7686

Qy   7872 TCAGTGACCAAGCTGCTTAGTGACCAACGAGTGACCGCTGGCAATTGCCGAGGCAGTTATACGCTGGGAGGAC 7931
Db   7687 TCAGTGACCAAGCTGCTTAGTGACCAACGAGTGACCGCTGGCAATTGCCGAGGCAGTATCACGCTGGGAGGAC 7746

Qy   7932 GTCTGTGTGTCGATGTGTCGATGTGGGTCACCGAGGCCACCATGCGATTATCGCTGCGATATCGCTCT 7991
Db   7747 GTCTGTGTGTCGATGTGTCGATGTGGGTCACCGAGGCCACCATGCGATATCGCGGGTCTT 7806

Qy   7992 TCAGTGACCAAGAGCTGCAACGAGTGACGCTGGCAATTGCCGAGGCAGTATACGCTGGGAGGAC 8051
Db   7807 TCAGTGACCAAGAGCTGCAACGAGTGACGCTGGCAATTGCCGAGGCAGTATCACGCTGGGAGGAC 7866

Qy   8052 CGATGATCGCTTATGATGGGCCGAGAGATCGGCGATATCGCCGGGTCT 8111
Db   7867 CGATGATCGCTTATGATGGGCCGAGAGATCGGCGATATCGGGGTCT 7926

Qy   8112 ATACTACCTTCAAGTTCCAACAGTTTGACCTGGCTGGCTGGAAGTGCTGGAAATGCTGCAGCCGAAC 8171

APPLICANT: JOHN N. SIMONS
APPLICANT: TAMI J. PILOT-MATIAS
APPLICANT: GEORGE J. DAWSON
APPLICANT: GEORGE G. SCHLAUDER
APPLICANT: SURESH M. DESAI
APPLICANT: THOMAS P. LEARY
APPLICANT: ANTHONY SCOTT MUERHOFF
APPLICANT: JAMES C. ERKER
APPLICANT: SHERI L. BUIJK
APPLICANT: ISA K. MUSHAHWAR

TITLE OF INVENTION: NON-A, NON-B. NON-C, NON-D, NON-E HEPATITIS
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
NUMBER OF SEQUENCES: 716

CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,550B
FILING DATE:
CLASSIFICATION: 435435

ATTORNEY/AGENT INFORMATION:
NAME: POREMBSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5527.PC.01

TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-938-2623

INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 4268 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-424-550B-80

Query Match   44.2%;   Score 4159;   DB 8;   Length 4268;
Best Local Similarity   98.1%;   Pred. No. 0;
Matches 4215;   Conservative 20;   Mismatches 30;   Indels 30;   Gaps 2;

```
Qy  3456 TGGCTCATCCACAGGCTCTATACACCCAATACCGTTGAGCGCGGCTAATGACCAGGACA 3515
         ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     1 TGGCTCATCCACAGGCTCTATACACCCAATAACCGTTGAGCGCGGCTAATGACCAGGACA 60

Qy  3516 TCTATCAACCACCATGTGGAGCTGGGGTCCCTTACTCGGTGTCCTTCGGTGTCCTTCAAGG 3575
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TCTATCAACCACCATGTGGAGCTGGGGTCCCTTACTCGGTGTCCTTCGGTGTCCTTCAAGG 120

Qy  3576 GGTATCTGGTAACGACTGGGGTCATTGGTGAGGTCAACAAATCCGATGACCCTTATT 3635
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGTATCTGGTAACGACTGGGGTCATTGGTGAGGTCAACAAATCCGATGACCCTTATT 180

Qy  3636 GGTGTGGTGTGCGGGGCCCTTCCCATGGCGTTGTTGCCAAGGGTTCTTCAGTGCCCCGATTC 3695
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GGTGTGGTGTGCGGGGCCCTTCCCATGGCGTTGTTGCCAAGGGTTCTTCAGTGCCCCGATTC 240

Qy  3696 TGTGCTCCCCGGGCATGTTATTGGGATGTTCACCGGTGTGCTGATACCCCCAGTACACAGCAC 3755
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGTGCTCCCCGGGCATGTTATTGGGATGTTCACCGGTGTGCTGATACCCCCAGTACACAGCAC 300

Qy  3756 TCAGTCAGATTAGGGTTAGGCCGTTGGTGTGTGCTGGATACCCCCAGTACACAGCAC 3815
         ||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TCGGCCAGATTAGGGTTAGGCCGTTGGTGTGTGCTGGATACCCCCAGTACACAGCAC 360
```

```
Db  7927 ATACTACCTCAAGTTCAACAGTTGACCTGCTGGCTGAAGGTAAATGCTGCACGCCCAAAC 7986
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       AGGCTGGCATGAAGAACCCTCGCTTCCTTATTTGCGGCGATGATTGCACCGTAATTTGGA 8231

Db  7987 AGGCTGGCATGAAGAACCCTCGCTTCCTTATTTGCGGCGATGATTGCACCGTAATTTGGA 8046
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       AGAGCGCCGGAGCAGATGCAGACAAACAAGCAATGCGTGTCTTGCTAGCTGGATGAAGG 8291

Db  8047 AGAGCGCCGGAGCAGATGCAGACAAACAAGCAATGCGTGTCTTGCTAGCTGGATGAAGG 8106
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       TGATGGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACACAGTTTGGAAGAATTAACAT 8351

Db  8107 TGATGGGTGCACCACAAGATTGTGTGCCTCAACCCAAATACACAGTTTGGAAGAATTAACAT 8166
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       CATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGCAAGCCTTACTTTCTTA 8411

Db  8167 CATGCTCATCAAATGTTACCTCTGGAATTACCAAAAGTGCAAGCCTTACTACTTTCTTA 8226
         ||||||||||||||||||||||||||||||||||||||||||||||||||: ||||||||
Qy       CAAGAGATCCTGTATCCCCCTTGGCAGGTGTCTGCCGAGGGTCTGGGATACAACCCCA 8471

Db  8227 CAAGAGATCCTCCTATCCCCCTTGGCAGGTGTCTGCCGAGGGTCTGGGATACAACCCCA 8286
         |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Qy       GTGCGTGCGGGATTGGGTATCTAATAACATCACTCACCCATGTTTGTGGCCGTGTGT 8531

Db  8287 GKGCKGGGTGGGATTGGGTATCTAATAACATCACTCACCCATGTTTGTGGCCGTGTGT 8346
         |:|:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       TGGCTGTGTCCATTTCATGGAGCAGAGTGCTCTTTGAGGACAAACTTCCCGAGACTGTGACCT 8591

Db  8347 TGGCTGTGTCCATTTCATGGAGCAGAGTGCTCTTTGAGGACAAACTTCCCGAGACTGTGACCT 8406
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       TTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGTGCCAGCCATCATTGCTG 8651

Db  8407 TTGACTGGTATGGGAAAAATTATACGGTGCCTGTAGAAGATCTGTGCCAGCCATCATTGCTG 8466
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       GTGTGCACGGTATTGAGGCTTTCTCGGTGCGCTACCAACGCTGAGATCCTCAGAG 8711

Db  8467 GTGTGCACGGTATTGAGGCTTTCTCGGTGCGCTACCAACGCTGAGATCCTCAGAG 8526
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       TTTCCCAATCACTCAACAGACCATGACCACTACGCCCCCCCTGCGAGCGAAAGAAAGCCA 8771

Db  8527 TTTCCCAATCACTCAACAGACCATGACCACTACGCCCCCCTGCGAGCGAAAGAAAGCCA 8586
         |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Qy       GGGCGGTCCTCCGCCAGCGCCAAGAGGGCCAAGAGGCGCGGCGAGCACGCGCAAATTGGCTCGCTTCC 8831

Db  8587 GGGCGGTCCTCCGCCAGCGCCAAGAGGCGGCGCCAAGAGGCGCGGCGAGCACG- AAAATTGGCTCGCTTCC 8645
         |||||||||||||||||||||||||||||| |||||||||||||||||||||||  ||||||||||||||||||
Qy       TTCTCTGCATGCTACATCTCTAGACCTCTACCAGATTTGGATAAGACGAGCCGTTGGCGTTGGCTCGGT 8891

Db  8646 TTCTCTGCATGCTACATCTCTAGACCTCTACCAGATTTGGATAAGACGAGCCGTTGGCGTTGGCTCGGT 8705
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       ACACCACTTTCAATTATTGTGATGTTACTCCCGGAGGGGGGATGTGTTATTACACCAC 8951

Db  8706 ACACCACTTTCAATTATTGTGATGTTACT-CCCGGAGGGGGGATGTGTTATTACACCAC 8764
         |||||||||||||||||||||||||||| :|||||||||||||||||||||||||||||
Qy       AGAGAAGATTGCAGAAGTTCCTTGTGAAGTTATTGGGCCTGTGCTTCATTGTTTTTGCCCTAGGGC 9011

Db  8765 AGAGAAGATTGCAGAAGTTCCTTGTGAAGTTATTGGGCCTGTGCTTCATTGTTTTTGCCCTAGGGC 8824
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       TCATTGCTGTGTTGGATTGGACTAGCCATCAGCTGAACCCCAAATTCAAAATTAACTAACAGTTTT 9071

Db  8825 TCATTGCTGTGTTGGATTGGACTAGCCATCAGCTGAACCCCAAATTCAAAATTAATTAACAGTTTT 8884
         ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Qy       TTTTTTTTTTTTTTTT 9090

Db  8885 TTTTTTTTTTTTTTTT 8903
         ||||||||||||||||
```

RESULT 6
US-08-424-550B-80
; Sequence 80, Application US/08424550B
; Publication No. US20020119447A1
; GENERAL INFORMATION:

```
Qy   3816 ATGCCACTCTTGATACAAACCTACTGTGCTAAGCAGTATTCAGTGCAAATTTTAATTG 3875
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    361 ATGCCACTCTTGATACAAAAACCTACTGTGCTAAGCAGTATTCAGTGCAAATTTTAATTG 420

Qy   3876 CCCCCACTGGCGCAAGTCAACCAAATTACCACTTCTTACATGCAGGAGAGTATG 3935
          |||||||||||||||||||||||||||||||||||||||||||||:|||:||||
Db    421 CCCCCACTGGCGCAGCGGCAAGTCAACCAAATTACCACTTCTTACATGCAGGRGAAGYATG 480

Qy   3936 AGGTCTTGGTCCTAAATCCCAGTGTGCTACAACAGCATCAAATGCCAAAGTACATGCACG 3995
          |||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db    481 AGGTCTTGGTCCTAAATCCCAGTGTGCTACAACAGCATCAAATGCCAAAGTACATGCACG 540

Qy   3996 CGACGTACGGCGTGAATCCAAATTGCTATTTAATGGCAAATGTACCAACACAGGGGCTT 4055
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CGACGTACGGCGTGAATCCAAATTGCTATTTAATGGCAAATGTACCAACACAGGGGCTT 600

Qy   4056 CACTTACGTACGCACACATATGGCATGTACCTGACCGGAGCATGTTCCCGGAACTATGATG 4115
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CACTTACGTACGCACACATATGGCATGTACCTGACCGGACGATGTTCCCGGAACTATGATG 660

Qy   4116 TAATCATTTGTGAGAATGCCATCCAGTGTGCTACAACCACCGTGTTGGGCATTGGAAAGG 4175
          ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db    661 TAATCATTTGTGACGAATGCCATCCAGTGTGCTACAACCACCGTGTTGGGCATTGGAAAGG 720

Qy   4176 TCCTAACCGAAGCTCCATCCAAAAATGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCC 4235
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TCCTAACCGAAGCTCCATCCAAAAATGTTAGGCTAGTGGTTCTTGCCACGGCTACCCCC 780

Qy   4236 CTGGGAGTAATCCCACCACCACCACCATGCCAACATGACAATAACTGAGATTCAATTAAC 4295
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CTGGGAGTAATCCCACCACCACCACCATGCCAACATGACAATAACTGAGATTCAATTAAC 840

Qy   4296 CTATCCCCTTTCATGGAAAAAAACTGTAGGAGAAAATCTGAAGAAAGGAGGACACCTTA 4355
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTATCCCCTTTCATGGAAAAAAACTGTAGGAGAAAATCTGAAGAAAGGAGGACACCTTA 900

Qy   4356 TCTTTGAGGTACCAAAAAAACTGTGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAA 4415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCTTTGAGGTACCAAAAAAACTGTGATGAGCTTGCTAACGAGTTAGCTCGAAAGGGAA 960

Qy   4416 TAACAGCTGTGTCTCTTACTACTATAGGGGATGTGACATCTCAAAAATCCCTGAGGCGTGTG 4475
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TAACAGCTGTGTCTCTTACTACTATAGGGGATGTGACATCTCAAAAATGCCTGAGGGGACTGTG 1020

Qy   4476 TAGTAGTTGCCACTGATGCCGTTGTGTACAGGTACACTGGTGACTTTGATTCCGTGTATG 4535
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TAGTAGTTGCCACTGATGCCGTTGTGTACAGGTACACTGGTGACTTTGATTCCGTGTATG 1080

Qy   4536 ACTGCAGCCTCATGTGTAGAGAGCGCAACATGCCATGCCTTGACCTGTTGACCCTACTTCACCATG 4595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ACTGCAGCCTCATGTGTAGAGAAGCCATGCCATGCCTTGACCTGTTGACCCTACTTCACCATG 1140

Qy   4596 GTGTTCGTGTGTGGGGTTCAGCAATGTTAAAGGCCAGCGTTGTACCCCGTAGGGGGCCGCACAGGCC 4655
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GTGTTCGTGTGTGGGGTTCAGCAATGTTAAAGGCCAGCGTTGTACCCCGTAGGGGGCCACAGGCC 1200

Qy   4656 GTGGGAGAGCTGGCATATACTACTATGTAGACGGGAGGTTGTACCCCTTCGGGTATGGTTC 4715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GTGGGAGAGCTGGCATATACTACTATGTAGACGGGAGGTTGTACCCCTTCGGGTATGGTTC 1260

Qy   4716 CTGAATGCAACATTGTTGAAGCTTGAGACGCACGCAGCCTTGAAGGCATGGTTTGTCATCAA 4775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CTGAATGCAACATTGTTGAAGCTTGAGACGCACGCAGCCTTGAAGGCATGGTTTGTCATCAA 1320

Qy   4776 CAGAAGCTCAAACTATTCTGGACACCTATCGCACCCAACCTGGGTTACCTTGCGATAGGAG 4835
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CAGAAGCTCAAACTATTCTGGACACCTATCGCACCCAACCTGGGTTACCTTGCGATAGGAG 1380

Qy   4836 CAAATTTGGACGAGTGGCTGCATCCTCTTTTCTATGGTCAACCCCGAACCTTCATTTGTCA 4895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CAAATTTGGACGAGTGGCTGCATCCTCTTTTCTATGGTCAACCCCGAACCTTCATTTGTCA 1440

Qy   4896 ATACTGCAAAAAGAAGAACTGCTCACAATTATGTTTTGTTGTTTTGTTGACTGCAGCCCAACTGT 4955
```

```
Db   1441 ATACTGCAAAAAGAAGAACTGCTGACAATTATGTTTTGTTGTTTTGTTGACTGCAGCCCAACTGT 1500

Qy   4956 GTCATCAGTATGGCTATGCTCCCAATGACGGCACCACCACCGGTGGCAGGCGGCGGCCTTG 5015
          |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db   1501 GTCATCAGTATGGCTATGCTCCCAATGACGGCACCACCACCGGTGGCAGGCGGCGGCCTTG 1560

Qy   5016 GGAAAAAACCTTGCTGTGGGGTTCGTGGCGGCCTTGACGGCGCTGTCCTGTCCTGGCCCAG 5075
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GGAAAAAACCTTGCTGTGGGGTTCGTGGCGGCCTTGACGGCGCCTGTCCTGTCCTGGCCCAG 1620

Qy   5076 AGCCCAGCGAGGTGACCAGATACCAAATGTGCTTCACTGAAGTCAATACTTCTGGGACAG 5135
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 AGCCCAGCGAGGTGACCAGATACCAAATGTGCTTCACTGAAGTCAATACTTCTGGGACAG 1680

Qy   5136 CCGCACTCGCTGCTGTGGCGCTGACACTTTGGCG 5195
          |||||:||||:||||||||||||||||||||||||
Db   1681 CCGCACTCGCTGCTGTATCTAGCTTATCTAGCACTTTTGGCG 1740

Qy   5196 CCACTTGTGTGCGGCGTTGCTGCTGCTGTCTATTACAATCAGTCCCTACCGGTGTCTGCGCCC 5255
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 CCACTTGTGTGCGGCGTTGCTGCTGCTGTCTATTACAATCAGTCCCTACCGGTGTCTGCGCCC 1800

Qy   5256 CAGTGGTTGACGAAGAGAAATCGTGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCA 5315
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CAGTGGTTGACGAAGAGAAATCGTGAGGAGTGTGCATCATTCATTCCCTTGGAGGCCA 1860

Qy   5316 TGGTTGCTGCAATTGACAAGCTGAAGAGTACAAATCACCACACAGTCCTTTCACATTGG 5375
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGGTTGCTGCAATTGACAAGCTGAAGAGTACAAATCACCACACAGTCCTTTCACATTGG 1920

Qy   5376 AAACCGGCCCTGAAAAAACTTAACACCTTTCTTGGGCCTCATGCCGCCACTGGGCTACTGCTA 5435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AAACCGGCCCTGAAAAAACTTAACACCTTTCTTGGGCCTCATGCCGCCACTGGCGCTACTGCTA 1980

Qy   5436 TCATAGAGTATTGCTGTGGTTTAGTCACTTTACCTGACAATCCCTTTGCATCATGCGTGT 5495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TCATAGAGTATTGCTGTGGCTTAGTCACTTTACCTGACAATCCCTTTGCATCATGCGTGT 2040

Qy   5496 TTGCTTTCATTGCGGGTATTACCCCACTACCTCACAAGATCAAAATGTTCCTGTCAT 5555
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TTGCTTTCATTGCGGGTATTACCCCACTACCTCACAAGATCAAAATGTTCCTGTCAT 2100

Qy   5556 TATTTGGAGGGCCAATTGCGTCCAAGTTGATTCCGTGGCACTGGCCACTGGCGTCATGA 5615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 TATTTGGAGGGCCAATTGCGTCCAAGTTGATTCCGTGGCACTGGCCACTGGCGTCATGA 2160

Qy   5616 TGGCCGGGCGTCGCGGAACAGCTGTTCTTGGTCATGACATGGCGGGTTTGTCTTGACA 5675
          |||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db   2161 TGGCCGGGGCGTCGYGGGAACAGCTGTTCTTGGTCATGACATGGCGGGTTTGTCTTGACA 2220

Qy   5676 TGCTAGGCGGCTGCATGCCGCCCTGCTTGCTTGACATTTAAATGCTTGATGG 5735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 TGCTAGGCGGCTGCATGCCGCCCTGCTTGCTTGACATTTAAATGCTTGATGG 2280

Qy   5736 GTGAGTGGCCCACTCAGCGATCAGTTCTACTCCGGTCAATCCGGTCCG 5795
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GTGAGTGGCYCACTCAGCGATCAGTTCTACTCCGGTCAATCCGGTCCG 2340

Qy   5796 CAGGAGTGTGTGGGCGTCTTGTCAGCTTGCAATGTTGCTTTGCAACAGCAGGAGGGGCCAG 5855
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CAGGAGTGTGTGGGCGTCTTGTCAGCTTGCAATGTTGCTTTGCAACAGCAGGAGGGGCCAG 2400

Qy   5856 ATCACTGGCCAACAGACTTCTTACTATGCTTGCTAGGAGCAACACTGTATGTAATGAGT 5915
          |||||||||||||||||||||:||||||||||||||||||||||||||||||:||||||
Db   2401 ATCACTGCCCAACAGACTTCTTACTATGCTTGCTAGGAGCAACACTGTATGTARTGAGT 2460

Qy   5916 ACTTTATTGCCACTCGTGACATCGCAGGAAGCATCTGACAGATCTGGGCATTCTACCCC 5975
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 ACTTTATTGCCACTCGTGACATCCGCAGGAAGCATCTGACAGATCTGGGCATTCTACCCC 2520

Qy   5976 CCTGGAGTGTCATATACGCTTGCACCCCGACGCGAGGATTGCGCG 6035
          ||:|||||||||||||||||||::|||||||||||||||||||||
Db   5976 CCTGGAGTGTCATATACGCTTGCACCCCGACGCGAGGATTGCGCG
```

```
Db  2521  CCTGGAGTRTCATATCAGCTTGCATCCGTTGGCTYACACCCGACGGAGGATGATTGCG  2580

Qy  6036  GCCTCATTGCTTGGGGTCTAGAGATTTGGGCAGTCAGATGTGTGCAATTTCTTTGTGATTGCT  6095
           ||||||||||||||:||||||||| ||||||||||||||||||||||||||||||||||||||
Db  2581  GCCTCATTGCTTGGGGTCTARAGATTTGGCAGTCAGATGTGTGCAATTTCTTTGTGATTTGCT  2640

Qy  6096  TTAATGTCCTTAAAGCTTGGAGTTCAGAGACATGGTTAACATTCCTGGTTGTCCTTTCTACA  6155
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  TTAATGTCCTTAAAGCTGGAGTTCAGAGACATGGTTAACATTCCTGGTTGTCCTTTCTACA  2700

Qy  6156  GCTGCCAGAAGGGGTACAAGGCCCCTGGATTGGATCAGGTATGCTCCAAGCACGTGTC  6215
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  GCTGCCGAGAAGGGGTACAAGGCCCCTGGATTGGATCAGGTATGCTCCAAGCACGTGTC  2760

Qy  6216  CATGGCGGTGCTGAACTCATCTTTCTGTGAGAATGGTTTTGCAAAACTTTACAAAGGAC  6275
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db  2761  CATGGCGGTGCTGAACTCATCTTTCTGTGAGAATGGTTTTGCAAAACTTTACAAAGGAC  2820

Qy  6276  CCAGAACTTGTTCAAATTACTGGAGAGGGCGTCTTCCAGTCAACGCTAGGCGTGTGTGGGT  6335
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  CCAGAACTTGTTCAAATTACTGGAGAGGGCGTCTTCCAGTCAACGCTAGGCGTGTGTGGGT  2880

Qy  6336  CGGCTAGACCGACCCAACTGATTGACTAGTCTTGTCGTCAATTATGGCGTTAGGGACT  6395
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  CGGCTAGACCGACCCAACTGATTGACTAGTCTTGTCGTCAATTATGGCGTTAGGGACT  2940

Qy  6396  ACTGTAAATATGAGAAAATGGGAGATCACATTTTGTTACAGCAGTATCCTCCAAATG  6455
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db  2941  ACTGTAAATATGAGAAATTGGGAGATCACATTTTGTTACAGCAGTATCCTCCAAATG  3000

Qy  6456  TCTGTTTCACCCAGGTGCCCCAACCTTGAGAGCTGCAGTGGCGTGGACGGCGTACAGG  6515
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  TCTGTTTCACCCAGGTGCCCCAACCTTGAGAGCTGCAGTGGCGTGGACGGCGTACAGG  3060

Qy  6516  TTCAGTGTTATCTAGGTGAGCCCAAAACTCCTTGGACGACATCGCTTGCTGTGTTACGGTC  6575
           ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db  3061  TTCAGYGTTATCTAGGTGAGCCCAAAACTCCTTGGACGACATCGCTTGCTGTGTTACGGTC  3120

Qy  6576  CTGGCGGTAAGGGTAAAACTGTTAAGCTTCCCCTCCGCGTTGCACGCGTCACACCTGGTG  6635
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  CTGACGGTAAGGGTAAAACTGTTAAGCTTCCCCTCCGCGTTGCACGCGTCACACCTGGTG  3180

Qy  6636  TGCGGCATGCAACTTAATTTGCGTGATGCACTTGAGACAATGACTGTAATTCCACAAACA  6695
           |||||||||||||||| |||||||||||||||||| |||||||||||||||||||||||
Db  3181  GTCGGCATGCAACTTAATTTGCGTGATGCACTTGAGGACAAATGACTGTAATTCCATAAACA  3240

Qy  6696  ACACTCCTAGTGATGAAGCCGCAGTGTCCGCTCTTGTTTTCAAACAGGAGTTGCGGGCTA  6755
           |||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db  3241  ACACTCCTAGTGATGAAGCCGCAGTGTCCGCTCTTGTTTTCAAACAGGAGTTGCGGGCTA  3300

Qy  6756  CAAACCAATTGCTTGAGGCAATTTCAGCTGGCGTTGACACCACCAAAACTGCCAGCCCCT  6815
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301  CAAACCAATTGCTTGAGGCAATTTCAGCTGGCGTTGACACCACCAAAACTGCCAGCCCCT  3360

Qy  6816  CC---ATCGGAGAGTAGTGGTAAGAAGCGCAGTTCCGGGCAAGAAACTGGTTCGCTTA  6872
           ||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361  CCCAGATCGAGAGTAGTGGTAAGAAGCGCAGTTCCGGGCAAGAAACTGGTTCGCTTA  3420

Qy  6873  CCTTGCCTCCCCTCCCGAGATCCGTCCCAGGAGTGTCATGTCCTGAAAGCTTGCAACGAA  6932
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421  CCTTGCCTCCCCTCCCGAGATCCGTCCCAGGAGTGTCATGTCCTGAAAGCCTGCAACGAA  3480

Qy  6933  GTGACCCGGTTAGAAGGTTCCTTCAAMCCTCCCTTCTTCACCACCTGTTCTRCAGTTGGCCA  6992
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481  GTGACCCGGTTAGAAGGTTCCTTCAAMCCTCCCTTCTTCACCACCTGTTCTRCAGTTGGCCA  3540

Qy  6993  TGCCGATGCCCCGTTGTGGGAGCCGGTGAGTGTAACCCTTTCACTGCAATTGGATGTGCAA  7052
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541  TGCCGATGCCCCGTTGTGGGAGCCGGTGAGTGTAACCCTTTCACTGCAATTGGATGTGCAA  3600

Qy  7053  TGACCGAAACAGGCGGAGGCCCTGATGATTTACCCAGTTACCCTGAAAAAGGAGAGTCT  7112
           ||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db  3601  TGACCGWAACARGCYGGGAGKCCCWSAKRATTTACCCAGTTACCCTCCCAAAAAGGAGAGTCT  3660
```

```
Qy  7113  CTGAATGTCAGACGAAAGTTGGTCGACGGCTACAACCGCTTCYCAGCTTACGTTACTGGCC  7172
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db  3661  CTGAATGGTCAGACGAAAGTTGGTCAACGACTACAACCCGTTCYCAGCTTACGTTACTGGCC  3720

Qy  7173  CCCCGTACCCTAAGATACGGGGAAAAGGATTCCACTCAGTGATTGTCAGCAGCCCCCCAAACGCGCCTA  7232
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721  CCCCGTACCCTAAGATACGGGGCAAGGATTCCACTCAATCAGCCACCGCCCAAACGCGCCTA  3780

Qy  7233  CAAAAAGAAGTTGGGAAAGAGTGAGTTTCGTGCAGCATGAGCTGGACCGACACG  7292
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781  CAAAAAGAAGTTGGGAAAGAGTGAGTTTCGTGCAGCATGAGCTGGACCGACACG  3840

Qy  7293  TGATTAGTTCAAAACTGCTTCTAAAGTTCTGTCTGCAACTCGGGCCATCACTAGTGGTT  7352
           |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db  3841  TGATTAGCTTCAAAACTGCTTCTAAAGTTCTGTCTGCAACTCGGGCCATCACTAGTGGTT  3900

Qy  7353  TCCTTCAAACAAAGATCATTGGTCATGTGTATGTGACTGCAGCGCGGGAGCTTAGAAAAC  7412
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901  TCCTCAAACAAAGATCATTGGTCATGTGTATGTGACTGCAGCGCGGGAGCTTAGAAAAC  3960

Qy  7413  AAAAAGTCACTATTAATAGACAACCTCGTTCCCCCATCATCACCACAAGCAAGTCAGACGAAGAT  7472
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961  AAAAAGTCACTATTAATAGACAACCTCTGTTCCCCCATCATCACCACAAGCAAGTCAGACGAAGAT  4020

Qy  7473  TGGCTAAGGAAAAAGCTTCAAAAGTTGTCGGTGTCATGTGGACTATGATGAAGTAGTAGCAG  7532
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021  TGGCTAAGGAAAAAGCTTCAAAAGTTGTCGGTGTCATGTGGACTATGATGAAGTAGTAGCAG  4080

Qy  7533  CTCACACGCCCCTCTAAGTCTGCTCAAGTCCACATCACTGCCTTCGGGGCACTCGATGTTC  7592
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081  CTCACACGCCCCTCTAAGTCTGCTCAAGTCCACATCACTGCCTTCGGGGCACTCGA-----  4135

Qy  7593  GTTCTGGAGCAGCCCGCAAGGCTGTTCTCGAGACTTGCAGAAGTGTGTCGAGGCAGGTGAGA  7652
           |||||||||||||||||||||||||||||     |||||||||||||||||||||||||||
Db  4136  --------------------------------TGTTCTGGACTTGCAGAAGTGTGTCGAGGCAGGTGAGA  4173

Qy  7653  TACCGAGTCATTATCGGCAAACTGTGATAGTTCCAAAGGAGAGGTCTTCGTGAAGACCC  7712
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4174  TACCGAGTCATTATCGGCAAACTGTGATAGTTCGTGATAGTTCGTGAAGAGGAGGTCTTCGTGAAGACCC  4233

Qy  7713  CCCAGAAACCAACAAAGAAACCCCAAGGCTTATC  7747
           |||||||||||||||||||||||||||||||||
Db  4234  CCCAGAAACCAACAAAGAAACCCCAAGGCTTATC  4268
```

RESULT 7
US-08-424-550B-25
; Sequence 25, Application US/08424550B
; Publication No. US20020119447A1
GENERAL INFORMATION:
 APPLICANT: JOHN N. SIMONS
 APPLICANT: TAMI J. PILOT-MATIAS
 APPLICANT: GEORGE J. DAWSON
 APPLICANT: GEORGE G. SCHLAUDER
 APPLICANT: SURESH M. DESAI
 APPLICANT: THOMAS P. LEARY
 APPLICANT: ANTHONY SCOTT MUERHOFF
 APPLICANT: JAMES C. ERKER
 APPLICANT: SHERI L. BUIJK
 APPLICANT: ISA K. MUSHAHWAR
 TITLE OF INVENTION: NON-A, NON-B. NON-C, NON-D, NON-E HEPATITIS
 TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
 NUMBER OF SEQUENCES: 716
 CORRESPONDENCE ADDRESS:
  ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
  STREET: 100 ABBOTT PARK ROAD
  CITY: ABBOTT PARK
  STATE: IL
  COUNTRY: USA
  ZIP: 60064-3500
 COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk

```
          COMPUTER: IBM PC compatible
          OPERATING SYSTEM: PC-DOS/MS-DOS
          SOFTWARE: PatentIn Release #1.0, Version #1.25
     CURRENT APPLICATION DATA:
          APPLICATION NUMBER: US/08/424,550B
          FILING DATE:
          CLASSIFICATION: 435435
     ATTORNEY/AGENT INFORMATION:
          NAME: POREMBSKI, PRISCILLA E.
          REGISTRATION NUMBER: 33,207
          REFERENCE/DOCKET NUMBER: 5527.PC.01
     TELECOMMUNICATION INFORMATION:
          TELEPHONE: 708-937-6365
          TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 25:
     SEQUENCE CHARACTERISTICS:
     LENGTH: 479 base pairs
     TYPE: nucleic acid
     STRANDEDNESS: single
     TOPOLOGY: linear
     MOLECULE TYPE: DNA (genomic)
US-08-424-550B-25

Query Match     4.8%;   Score 449.8;  DB 8;  Length 479;
Best Local Similarity  97.5%;   Pred. No. 2.8e-121;
Matches  467;  Conservative   1;  Mismatches   8;  Indels   3;  Gaps   1;

Qy  6419  GATCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCA  6478
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  GATCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCA  60

Qy  6479  ACCTTGAGAGCTGCAGTGGCCGTGGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCC  6538
                |||||||||||||||||||||||||||||||| ::|||||||||||||||||||||
Db    61  ACCTTGAGAGCTGCAGTGGCCGTGGACGGCGTACAGGTTCAGYGTTATCTAGGTGAGCCC  120

Qy  6539  AAAACTCCTTGGACGACATCTGCTTGCTGTTACGGTCCTGACGGTAAGGTAAAACTGTT  6598
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  AAAACTCCTTGGACGACATCTGCTTGCTGTTACGGTCCTGACGGTAAGGTAAAACTGTT  180

Qy  6599  AAGCTTCCCTTCCGCGGTGACGGTGTGCGCATGCAACTGGTCGCCATGCAACTTAATTGCGT  6658
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  AAGCTTCCCTTCCGCGGTGACGGTGTGCGCATGCAACTGGTCGCCATGCAACTTAATTGCGT  240

Qy  6659  GATGCACTTGAGACAAATGACTGTAATTCCACAAACAACACTCCTAGTGATGAAGCCGCA  6718
                |||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   241  GATGCACTTGAGACAAATGACTGTAATTCCATAAACAACACTCCTAGTGATGAAGCCGCA  300

Qy  6719  GTGTCCGCTCTTGTTTCAAACAGGAGTTGCGGCTACAACCAATTGCTTGAGGCAATT  6778
                |||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   301  GTGTCCGCTCTTGTTTCAAACAGGAGTTGCGGCTACAAACCAATTGCTTGAGGCAATT  360

Qy  6779  TCAGCTGGCGTTGACACCACCACCAAACTGCCAGCCCCTCC---ATCGAAGAGGTAGTGGTA  6835
                ||||||||||||||||||||||||||||||||||||||   |||||||||||||||||
Db   361  TCAGCTGGCGTTGACACCACCACCAAACTGCCAGCCCCCTCCCCGATCGAAGAGGTAGTGGTA  420

Qy  6836  AGAAAGCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCCCTCCGAGATC  6894
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  AGAAAGCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCCCTCCGAGATC  479
```

          APPLICANT: SHERI L. BUIJK
          APPLICANT: ISA K. MUSHAHWAR
          TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
          TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
          NUMBER OF SEQUENCES: 716
          CORRESPONDENCE ADDRESS:
          ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
          STREET: 100 ABBOTT PARK ROAD
          CITY: ABBOTT PARK
          STATE: IL
          COUNTRY: USA
          ZIP: 60064-3500
```
     COMPUTER READABLE FORM:
          MEDIUM TYPE: Floppy disk
          COMPUTER: IBM PC compatible
          OPERATING SYSTEM: PC-DOS/MS-DOS
          SOFTWARE: PatentIn Release #1.0, Version #1.25
     CURRENT APPLICATION DATA:
          APPLICATION NUMBER: US/08/424,550B
          FILING DATE:
          CLASSIFICATION: 435435
     ATTORNEY/AGENT INFORMATION:
          NAME: POREMBSKI, PRISCILLA E.
          REGISTRATION NUMBER: 33,207
          REFERENCE/DOCKET NUMBER: 5527.PC.01
     TELECOMMUNICATION INFORMATION:
          TELEPHONE: 708-937-6365
          TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 392:
     SEQUENCE CHARACTERISTICS:
     LENGTH: 479 base pairs
     TYPE: nucleic acid
     STRANDEDNESS: single
     TOPOLOGY: linear
     MOLECULE TYPE: DNA (genomic)
US-08-424-550B-392

Query Match     4.8%;   Score 449.8;  DB 8;  Length 479;
Best Local Similarity  97.5%;   Pred. No. 2.8e-121;
Matches  467;  Conservative   1;  Mismatches   8;  Indels   8;  Gaps   1;

Qy  6419  GATCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCA  6478
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  GATCACATTTTGTTACAGCAGTATCCTCTCCAAATGTCTGTTTCACCCAGGTGCCCCA  60

Qy  6479  ACCTTGAGAGCTGCAGTGGCCGTGGACGGCGTACAGGTTCAGTGTTATCTAGGTGAGCCC  6538
                ||||||||||||||||||||||||||||||||| ::|||||||||||||||||||||
Db    61  ACCTTGAGAGCTGCAGTGGCCGTGGACGGCGTACAGGTTCAGYGTTATCTAGGTGAGCCC  120

Qy  6539  AAAACTCCTTGGACGACATCTGCTTGCTGTTACGGTCCTGACGGTAAGGTAAAACTGTT  6598
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  AAAACTCCTTGGACGACATCTGCTTGCTGTTACGGTCCTGACGGTAAGGTAAAACTGTT  180

Qy  6599  AAGCTTCCCTTCCGCGGTGACGGTGTGCGCATGCAACTGGTCGCCATGCAACTTAATTGCGT  6658
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  AAGCTTCCCTTCCGCGGTGACGGTGTGCGCATGCAACTGGTCGCCATGCAACTTAATTGCGT  240

Qy  6659  GATGCACTTGAGACAAATGACTGTAATTCCACAAACAACACTCCTAGTGATGAAGCCGCA  6718
                |||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   241  GATGCACTTGAGACAAATGACTGTAATTCCATAAACAACACTCCTAGTGATGAAGCCGCA  300

Qy  6719  GTGTCCGCTCTTGTTTCAAACAGGAGTTGCGGCTACAACCAATTGCTTGAGGCAATT  6778
                |||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   301  GTGTCCGCTCTTGTTTCAAACAGGAGTTGCGGCTACAAACCAATTGCTTGAGGCAATT  360

Qy  6779  TCAGCTGGCGTTGACACCACCACCAAACTGCCAGCCCCTCC---ATCGAAGAGGTAGTGGTA  6835
                ||||||||||||||||||||||||||||||||||||||   |||||||||||||||||
Db   361  TCAGCTGGCGTTGACACCACCACCAAACTGCCAGCCCCCTCCCCGATCGAAGAGGTAGTGGTA  420

Qy  6836  AGAAAGCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCCCTCCGAGATC  6894
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  AGAAAGCGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTTGCCTCCCCCCTCCGAGATC  479
```

RESULT 8
US-08-424-550B-392
; Sequence 392, Application US/08424550B
; Publication No. US20020119447A1
; GENERAL INFORMATION:
;   APPLICANT: JOHN N. SIMONS
;   APPLICANT: TAMI J. PILOT-MATIAS
;   APPLICANT: GEORGE J. DAWSON
;   APPLICANT: GEORGE G. SCHLAUDER
;   APPLICANT: SURESH M. DESAI
;   APPLICANT: THOMAS P. LEARY
;   APPLICANT: ANTHONY SCOTT MUERHOFF
;   APPLICANT: JAMES C. ERKER

RESULT 9
US-08-424-550B-29/c
; Sequence 29, Application US/08424550B
; Publication No. US20020119447A1
GENERAL INFORMATION:
; ;   APPLICANT: JOHN N. SIMONS
; ;   APPLICANT: TAMI J. PILOT-MATIAS
; ;   APPLICANT: GEORGE J. DAWSON
; ;   APPLICANT: GEORGE G. SCHLAUDER
; ;   APPLICANT: SURESH M. DESAI
; ;   APPLICANT: THOMAS P. LEARY
; ;   APPLICANT: ANTHONY SCOTT MUERHOFF
; ;   APPLICANT: JAMES C. ERKER
; ;   APPLICANT: SHERI L. BUIJK
; ;   APPLICANT: ISA K. MUSHAHWAR
; ;   TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; ;   TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
; ;   NUMBER OF SEQUENCES: 716
; ;   CORRESPONDENCE ADDRESS:
; ;   ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; ;   STREET: 100 ABBOTT PARK ROAD
; ;   CITY: ABBOTT PARK
; ;   STATE: IL
; ;   COUNTRY: USA
; ;   ZIP: 60064-3500
; ;   COMPUTER READABLE FORM:
; ;   MEDIUM TYPE: Floppy disk
; ;   COMPUTER: IBM PC compatible
; ;   OPERATING SYSTEM: PC-DOS/MS-DOS
; ;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; ;   CURRENT APPLICATION DATA:
; ;   APPLICATION NUMBER: US/08/424,550B
; ;   FILING DATE:
; ;   CLASSIFICATION: 435435
; ;   ATTORNEY/AGENT INFORMATION:
; ;   NAME: POREMBSKI, PRISCILLA E.
; ;   REGISTRATION NUMBER: 33,207
; ;   REFERENCE/DOCKET NUMBER: 5527.PC.01
; ;   TELECOMMUNICATION INFORMATION:
; ;   TELEPHONE: 708-937-6365
; ;   TELEFAX: 708-938-2623
; ; INFORMATION FOR SEQ ID NO: 29:
; ;   SEQUENCE CHARACTERISTICS:
; ;   LENGTH: 337 base pairs
; ;   TYPE: nucleic acid
; ;   STRANDEDNESS: single
; ;   TOPOLOGY: linear
; ;   MOLECULE TYPE: DNA (genomic)
US-08-424-550B-29

Query Match   3.4%;   Score 319.6; DB 8; Length 337;
Best Local Similarity   98.5%; Pred. No. 7.2e-83;
Matches 333; Conservative   0; Mismatches   4;   Indels   1; Gaps   1;

Qy   5855 GATCACTGGCCCAACAGACTTCTTACTATGCTTGCTAGGAGCAACACTGTATGTAATGAG 5914
           ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   337 GATCACTGCCCAACAGACTTCTTACTATGCTTGCTAGGAGCAACACTGTATGTAGTGAG   278

Qy   5915 TACTTTATTGCCACTCGTGACATCCGCAGGAGAGATACTGGGCATTCTGGAGGCATCTACC 5974
           ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   277 TACTTTATTGCCACTCGTGACATCCGCA-GAAGATACTGGGCATTCTGGAGGCATCTACC   219

Qy   5975 CCCTGGAGTGTCATATCAGTTGGCAGTATGTGTGCAATTTCTTTGTGATTTGC 6034
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   218 CCCTGGAGTGTCATATCAGTTGGCAGTATGTGTGCAATTTCTTTGTGATTTGC   159

Qy   6035 GGCCTCATTGCTTGGGGTCTAGAGATTTGGCAGTATGTGTGCAATTTCTTTGTGATTTGC 6094
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   158 GGCCTCATTGCTTGGGGTCTAGAGATTTGGCAGTATGTGTGCAATTTCTTTGTGATTTGC   99

Qy   6095 TTTAATGTCCTTAAAGCTGGAGTTCAGAGACATGGTTAACATTCCTGGTTGTCCTTTCTAC 6154
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   98 TTTAATGTCCTTAAAGCTGGAGTTCAGAGCCCCTGGATTGGATC 6192
           |||||||||||||||||||||||||
Qy   6155 AGCTGCCAGAAGGGGTACAAGGGCCCTGGATTGGATC 6192
           |||||||||||||||||||||||||||||||||||||
Db   38 AGCTGCCAGGAGGAGGGTACGAGGGGCCCTGGATTGGATC   1

RESULT 10
US-08-424-550B-159/c
; Sequence 159, Application US/08424550B
; Publication No. US20020119447A1
GENERAL INFORMATION:
; ;   APPLICANT: JOHN N. SIMONS
; ;   APPLICANT: TAMI J. PILOT-MATIAS
; ;   APPLICANT: GEORGE J. DAWSON
; ;   APPLICANT: GEORGE G. SCHLAUDER
; ;   APPLICANT: SURESH M. DESAI
; ;   APPLICANT: THOMAS P. LEARY
; ;   APPLICANT: ANTHONY SCOTT MUERHOFF
; ;   APPLICANT: JAMES C. ERKER
; ;   APPLICANT: SHERI L. BUIJK
; ;   APPLICANT: ISA K. MUSHAHWAR
; ;   TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; ;   TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
; ;   NUMBER OF SEQUENCES: 716
; ;   CORRESPONDENCE ADDRESS:
; ;   ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; ;   STREET: 100 ABBOTT PARK ROAD
; ;   CITY: ABBOTT PARK
; ;   STATE: IL
; ;   COUNTRY: USA
; ;   ZIP: 60064-3500
; ;   COMPUTER READABLE FORM:
; ;   MEDIUM TYPE: Floppy disk
; ;   COMPUTER: IBM PC compatible
; ;   OPERATING SYSTEM: PC-DOS/MS-DOS
; ;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; ;   CURRENT APPLICATION DATA:
; ;   APPLICATION NUMBER: US/08/424,550B
; ;   FILING DATE:
; ;   CLASSIFICATION: 435435
; ;   ATTORNEY/AGENT INFORMATION:
; ;   NAME: POREMBSKI, PRISCILLA E.
; ;   REGISTRATION NUMBER: 33,207
; ;   REFERENCE/DOCKET NUMBER: 5527.PC.01
; ;   TELECOMMUNICATION INFORMATION:
; ;   TELEPHONE: 708-937-6365
; ;   TELEFAX: 708-938-2623
; ; INFORMATION FOR SEQ ID NO: 159:
; ;   SEQUENCE CHARACTERISTICS:
; ;   LENGTH: 337 base pairs
; ;   TYPE: nucleic acid
; ;   STRANDEDNESS: single
; ;   TOPOLOGY: linear
; ;   MOLECULE TYPE: DNA (genomic)
US-08-424-550B-159

Query Match   3.4%;   Score 319.6; DB 8; Length 337;
Best Local Similarity   98.5%; Pred. No. 7.2e-83;
Matches 333; Conservative   0; Mismatches   4;   Indels   1; Gaps   1;

Qy   5855 GATCACTGGCCCAACAGACTTCTTACTATGCTTGCTAGGAGCAACACTGTATGTAATGAG 5914
           ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   337 GATCACTGCCCAACAGACTTCTTACTATGCTTGCTAGGAGCAACACTGTATGTAGTGAG   278

Qy   5915 TACTTTATTGCCACTCGTGACATCCGTGACATCCGCAGGAGAAGATACTGGGCATCTACC 5974
           ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   337 TACTTTATTGCCACTCGTGACATCCGTGACATCCGCA-GAAGATACTGGGCATCTACC   219

Qy   5915 TACTTTATTGCCACTCGTGACATCAGCTTGCATCCGCAGGAGAAGATACTGGGCATCTACC 5974
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   277 TACTTTATTGCCACTCGTGACATCAGCTTGCATCCGCA--GAAGATACTGGGCATCTACC   219

Qy   5975 CCCTGGAGTGTCATATGATCAGCTTGCATATCAGCTTGCCACACCCCGACGGAGGATTGC 6034
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   218 CCCTGGAGTGTCATATGATCATCCGCATATCAGCTTGCCACACCCCGACGGAGGATTGC   159

Qy   5975 TACTTTATTGCCACTCGTGACATCCGCAGGAGAAGATACTGGGCATCTACC 6034
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   277 TACTTTATTGCCACTCGTGACATCCGCA-GAAGATACTGGGCATCTACC   219

Qy   5975 CCCTGGAGTGTCATATCAGTTGGCAGTATGTGTGCAATTTCTTTGTGATTGC 6034
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   277 CCCTGGAGTGTCATATCAGTTGGCAGTATGTGTGCAA-GAAGATACTGGGCATCTACC   219

Qy   5975 CCCTGGAGTGTCATATCAGTTGGCAGTATGTGTGCAATTTCTTTGTGATTGC 6034
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   218 CCCTGGAGTGTCATATAATCAGCTTGCATCATCCTTGGCAGTATGTGTGCAATTTCTTTGTGATTGC   159

Qy   5855 GATCACTGCCGCCCAACAGACTTCTTACTATGCTTGCTAGGAGCAACACTGTATGTAGTGAG 5914
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   337 GATCACTGCCGCCCAACAGACTTCTTACTATGCTTGCTAGGAGCAACACTGTATGTAGTGAG   278

Qy   5915 TACTTAATTGCCACTCGTGACATCGTGACATCCTTACTATGCTTACTATGCTTGGAGGCATCTACC 5974
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   337 GATCACTGCCCAACAGACTTCTTACTATGCTTACTATGCTTGCTAGGAGCAACACTGTATGTAGTGAG   278

Qy   5915 TACTTTATTGCCACTCGTGACATCCTGACATCCGCAGGAGAAGATACTGGGCATCTACC 5974
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   277 TACTTTATTGCCACTCGTGACATCCGTGACATCCGCA-GAAGATACTGGGCATCTACC   219

Qy   5975 CCCTGGAGTGTCATATCAGTTGGCAGTATGTGTGCAATTTCTTTGTGATTGC 6034
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   277 CCCTGGAGTGTCATATCAGTTGGCAGTATGTGTGCAATTTCTTTGTGATTGC   219

Qy   98 TTTAATGTCCTTAAAGCTGGAGTTCAGAGCATGGTTAACATTCCTGGTTGTCCTTTCTAC   39
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6155 AGCTGCCAGAAGGGGTACAAGGGCCCTGGATTGGATC 6192
           |||||||||||||||||||||||||
Qy   38 AGCTGCCAGGAGGAGGGTACGAGGGGCCCTGGATTGGATC   1

```
Qy  6035 GGCCTCATTGCTTGGGGTCTAGAGATTGGCAGTATGTGTGCAATTTCTTTGTGATTTGC 6094
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   158 GGCCTCATTGCTTGGGGTCTAGAGATTGGCAGTATGTGTGCAATTTCTTTGTGATTTGC 99

Qy  6095 TTTAATGTCCTTAAAGCTGGAGTTCAGAGACATGGTAACATTCCTGGTTGTCCTTTCTAC 6154
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    98 TTTAATGTCCTTAAAGCTGGAGTTCAGAGACATGGTAACATTCCTGGTTGTCCTTTCTAC 39

Qy  6155 AGCTGCCAGAAGGGGTACAAGGGCCCCTGGATTGGATC 6192
         ||||||||||||||||||||||||||||||||||||||
Db    38 AGCTGCCAGGAGGGGTACGAGGGCCCCTGGATTGGATC 1
```

RESULT 11
US-08-424-550B-18

```
; Sequence 18, Application US/08424550B
; Publication No. US20020119447A1
; GENERAL INFORMATION:
;    APPLICANT: JOHN N. SIMONS
;    APPLICANT: TAMI J. PILOT-MATIAS
;    APPLICANT: GEORGE J. DAWSON
;    APPLICANT: GEORGE G. SCHLAUDER
;    APPLICANT: SURESH M. DESAI
;    APPLICANT: THOMAS P. LEARY
;    APPLICANT: ANTHONY SCOTT MUERHOFF
;    APPLICANT: JAMES C. ERKER
;    APPLICANT: SHERI L. BUIJK
;    APPLICANT: ISA K. MUSHAHWAR
;    TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
;    TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
;    NUMBER OF SEQUENCES: 716
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
;       STREET: 100 ABBOTT PARK ROAD
;       CITY: ABBOTT PARK
;       STATE: IL
;       COUNTRY: USA
;       ZIP: 60064-3500
;    COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/424,550B
;       FILING DATE:
;       CLASSIFICATION: 435435
;    ATTORNEY/AGENT INFORMATION:
;       NAME: POREMBSKI, PRISCILLA E.
;       REGISTRATION NUMBER: 33,207
;       REFERENCE/DOCKET NUMBER: 5527.PC.01
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 708-937-6365
;       TELEFAX: 708-938-2623
;    INFORMATION FOR SEQ ID NO: 18:
;       SEQUENCE CHARACTERISTICS:
;          LENGTH: 281 base pairs
;          TYPE: nucleic acid
;          STRANDEDNESS: single
;          TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
US-08-424-550B-18
```

```
Query Match             3.0%;  Score 279.4;  DB 8;  Length 281;
Best Local Similarity  99.6%;  Pred. No. 5e-71;
Matches  280;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy  6637 GCGCATGCAACTTAATTTGCGTGATGCACTTGAGACAAATGACTGTAATTCCACAAACAA 6696
         ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     1 GCGCATGCAACTTAATTTGCGTGATGCACTTGAGACAAATGACTGTAATTCCATAAACAA 60

Qy  6697 CACTCCTAGTGATGAAGCCGCAGTGTCCGCTCCTTGTTTCAAACAGGAGTTGCGGCGTAC 6756
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    61 CACTCCTAGTGATGAAGCCGCAGTGTCCGCTCCTTGTTTCAAACAGGAGTTGCGGCGTAC 120

Qy  6757 AAACCAATTGCTTGAGGCAATTCAGCTGGCGGTTGACACCACCAAACTGCCAGCCCCTC 6816
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AAACCAATTGCTTGAGGCAATTCAGCTGGCGGTTGACACCACCAAACTGCCAGCCCCCTC 180

Qy  6817 CATCGAAGAGGTAGTGGTAAGAAAGCGGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTT 6876
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CATCGAAGAGGTAGTGGTAAGAAAGCGGCCAGTTCCGGGCAAGAACTGGTTCGCTTACCTT 240

Qy  6877 GCCTCCCCTCCGAGATCCGTCCCAGAGTGTCATGTCCTG 6917
         |||||||||||||||||||||||||||||||||||||||
Db   241 GCCTCCCCTCCGAGATCCGTCCCAGAGTGTCATGTCCTG 281
```

RESULT 12
US-08-424-550B-20

```
; Sequence 20, Application US/08424550B
; Publication No. US20020119447A1
; GENERAL INFORMATION:
;    APPLICANT: JOHN N. SIMONS
;    APPLICANT: TAMI J. PILOT-MATIAS
;    APPLICANT: GEORGE J. DAWSON
;    APPLICANT: GEORGE G. SCHLAUDER
;    APPLICANT: SURESH M. DESAI
;    APPLICANT: THOMAS P. LEARY
;    APPLICANT: ANTHONY SCOTT MUERHOFF
;    APPLICANT: JAMES C. ERKER
;    APPLICANT: SHERI L. BUIJK
;    APPLICANT: ISA K. MUSHAHWAR
;    TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
;    TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
;    NUMBER OF SEQUENCES: 716
;    CORRESPONDENCE ADDRESS:
;       ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
;       STREET: 100 ABBOTT PARK ROAD
;       CITY: ABBOTT PARK
;       STATE: IL
;       COUNTRY: USA
;       ZIP: 60064-3500
;    COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/424,550B
;       FILING DATE:
;       CLASSIFICATION: 435435
;    ATTORNEY/AGENT INFORMATION:
;       NAME: POREMBSKI, PRISCILLA E.
;       REGISTRATION NUMBER: 33,207
;       REFERENCE/DOCKET NUMBER: 5527.PC.01
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 708-937-6365
;       TELEFAX: 708-938-2623
;    INFORMATION FOR SEQ ID NO: 20:
;       SEQUENCE CHARACTERISTICS:
;          LENGTH: 281 base pairs
;          TYPE: nucleic acid
;          STRANDEDNESS: single
;          TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
US-08-424-550B-20
```

```
Query Match             3.0%;  Score 279.4;  DB 8;  Length 281;
Best Local Similarity  99.6%;  Pred. No. 5e-71;
Matches  280;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy  6637 GCGCATGCAACTTAATTTGCGTGATGCACTTGAGACAAATGACTGTAATTCCACAAACAA 6696
         ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     1 GCGCATGCAACTTAATTTGCGTGATGCACTTGAGACAAATGACTGTAATTCCATAAACA 60
```

```
Qy   6697 CACTCCTAGTGATGAAGCCGCAGTGTCCGCTCTTGTTTTCAAACAGGAGTTGCGGCGTAC 6756
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     61 CACTCCTAGTGATGAAGCCGCAGTGTCCGCTCTTGTTTTTCAAACAGGAGTTGCGGCGTAC 120

Qy   6757 AAACCAATTGCTTGAGGCAATTTCAGCTGGCGTTGGCGTTGACACCACCAAACTGCCAGCCCCCTC 6816
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AAACCAATTGCTTGAGGCAATTTCAGCTGGCGTTGGCGTTGACACCACCAAACTGCCAGCCCCCTC 180

Qy   6817 CATCGAAGAGGTAGTGTAAGAAAGGCCGCCAGTTCCGGGCAAGAACTGGTTCCGCTTACCTT 6876
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CATCGAAGAGGTAGTGTAAGAAAGGCCGCCAGTTCCGGGCAAGAACTGGTTCCGCTTACCTT 240

Qy   6877 GCCTCCCCCTCCGGAATCCGTCCCAGGAGTGTCATGTCCTG 6917
          ||||||||||||||||||||||||||||||||||||||||
Db    241 GCCTCCCCCTCCGGAATCCGTCCCAGGAGTGTCATGTCCTG 281
```

RESULT 13
US-10-328-127-1
US-10-328-127-1

Query Match 2.8%; Score 262.4; DB 17; Length 5211;
Best Local Similarity 49.6%; Pred. No. 2.8e-65;
Matches 733; Conservative 0; Mismatches 736; Indels 9; Gaps 2;

```
Qy   3411 ACAACGTGTTGTATACTGCTCACCATGGCAGCAAGGGGCGCCGGTTGGCTCATCCCACAG 3470
          |||||||||  |||||  ||  ||||||||||||||||||||||||||||||||||||||
Db    155 ACGGCGGTGTGTTGGACTGTCTTCCATGGCGTCTTCCATGGCCGGCTCAAAGACCTTGGCGGCCCCAAAG 214

Qy   3471 GCTCTATACACCCAATAACCGTTGACGCGGCGTAATGCGGCGTAATCATCAACCACCAT 3530
          |||||||  ||| |||||  ||| |||||||||||||  ||||||||||||||||||||
Db    215 GCCCGATCACCCAGATGTACACTAATGTGACCAGGCACCTCGTCGGCCTGGCAGGCGCCC 274

Qy   3531 GTGGAGCTGGGTCCCTTACTCGGTGCTCCTTGGCGGGAGACCAAGGGTATCGGTAACAC 3590
          ||| ||||||  |||  |||||||||  |||||||||||||||||||||||||||||||
Db    275 CTGGGGGGCGCGCTCCATGACACCATGCCAGCCAGCTCGACCGTCGCGCTCGCAGCGGAC 334

Qy   3591 GACTGGGGTCATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGCGGGG 3650
          |||||||||||||||||||||||  |||||||||||  ||||||||||||||||||||
Db    335 GACATGCCAGCGTCATTCCGGTGCCGCCCGGGGGCGCGAGTCAGAGTGAGGGGGAGCCCTGCTCTCCC 394

Qy   3651 CCCTTGCCATGCGTGTTGCCAAGGGTTCTTGCGCGTCTAGAAATTCGGCGGTTCAGTAGG 3710
          |||||||||||  |||  |||||||||||  ||| ||||||||||||||||||||||||
Db    395 CCAGGCCGTGCTCTCCTACTTGAAGGGCGTCTGCTTCCCGCTCTCTGCCCCTTCGGGGGC 454

Qy   3711 ATGTTATTGGGATGTTCACCGCTGCTGCTAGATACCATCCCCAGTACACAGCACATGCCACTCTTGATA 3770
          ||||| |||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db    455 ACGCGTGTGGGCATCTTCCGGGTCTGTGTGCACCCGGGGGGTTGCAAAAGCCGGTGGACT 514

Qy   3771 TTAGGCCGTTGGTGTGTGCTGATACCATCCCCAGTACACAGCACATGCCACTCTTGATA 3830
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    515 TCATACCTGTTGAGTCTATGGAAACTACTATGCGGTCTCCGGTCTTCACAGACAACTCAT 574
```

```
Qy   3831 CAAAACCTACTGTGCCTAACGAGTATTCAGTGCAAATTTAATTGCCCCCACTGGCAGCG 3890
          |||||||||||||||||||||||||||||||||||||||
Db    575 CCCCCCAGCCGTACCGCCAGCATTCCAAGTGGCCCATCTCACACGTCCTCTACTGCGAGCG 634

Qy   3891 GCAAGTCAACCAAATTACCACTTTCTTACATGCAGCGGAGAGTATGAGGTCTTGGTCCTAA 3950
          |||||||| ||| |||  || |||||||
Db    635 GCAAGAGCACTAAAGTGCCGGCTGCTTATGCAGCGTTGACACCGTACTTGTTCTTGA 694

Qy   3951 ATCCCAGTGTGGCTACAACAGCATCAATGCCAAAGTACATGCACGGCGACGTACGGCGTGA 4010
          ||  ||| ||  |||||||||||||||||||||||||
Db    695 ACCGTCCGTCGCGTTAGGTTTTGGGGGCGTATATGTCAAAGGACGTGCACCG 754

Qy   4011 ATCCAAATTGCTATTTTAATGGCAAATGTACCAACACAGGGGCTTCACTTACGTACAGCA 4070
          |||| |||| ||  ||
Db    755 ACCTAACATCAGAACTGGGGTAAGGACCATCACCACCACGAGGCGCCCCATCACGTACTCCA 814

Qy   4071 CATATGGCATGTACCTGACCGGAGCATGTGCCC------GGAACTATGATGTAATCATT 4124
          || |||||||||  |||| ||| || |||
Db    815 CCTATGGCAAGTTCCTTGCCGACGGTGTTGCTCTGGGGCATTGGGAAAGGTCCTAACCG 874

Qy   4125 GTGACGAATGCCATGCTACCGACGCAACCACCGTGTTGGGCATTGGGAAAGTCCTCCGACC 4184
          |||||||||||  |||||
Db    875 GTGACGAGTGCCACTCAACTGACTCGACTACCATCTTGGCGCATGCCACCAGTCCTGGACC 934

Qy   4185 AAGCTCCATCCAAAAAATGTTAGGCTAGTTGTTCTTGCCACGGTACCCCCCCTGGAGTAA 4244
          || ||| |||||||||||  |||||
Db    935 AAGGGGAGACGGGCTGGAGCGCGGCCTTGTCGTGCGTCCGCCACCGCTACGCCTCCGGGATCGG 994

Qy   4245 TCCCTACCACCATGCCAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCT 4304
          ||| ||||  ||||||||
Db    995 TCACGGTGCCACACCCCAATATGCAGGAGGTGCCCTGTCCAACACTGGAGAGATCCCCT 1054

Qy   4305 TTCATGGAAAAGATTAAGGAGGAAAAATCTGAAGAAAGGGAGGCACGTTATCTTGAGG 4364
          || ||||  |||| |||||
Db   1055 TCTATGGCAAGGCCATCCCATGCGAGGTCATCAAGGGGGAAGGGCATCTCCATTTTCTGCC 1114

Qy   4365 CTACCAAAAACACTGTGATGAGCTTGCTCAACGAGTTAGCTCGAAAGGAATAACAGCTG 4424
          || ||||||| |||||
Db   1115 ATTCCAAAAAGAGAGTGTGATGAGCTTGAGCTCGTCGGGCCTCGGACTCAACGCCG 1174

Qy   4425 TCTCTTACTATAGGGATGGATGTGACATCTCAAAAATCC---CTGAGGGGACTGTGTAGTAG 4481
          |||  |||||||||
Db   1175 TAGCGTATTACCGGGGCCTTGTGTCCGTCATACCGACCAGCGGAGACGTCGGTTGTTG 1234

Qy   4482 TTGGCACTGATGCCTTGTGTACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTGCA 4541
          ||| ||||||||  |||||
Db   1235 TGGCAACAGACGCCTTAATGACGGCCTTTACCGGCGACTTTGATTCAGTGATGCACTGCA 1294

Qy   4542 GCCTCATGGTAGAAGGCACATGCCATTGACCTTGACCCTACTTTCACCATGGGTGTTC 4601
          |||  |||||||||
Db   1295 ATACGTGTTACCCAGACAGTGATTCAGTCGGGAGCAGTTCACCATTGAGACGA 1354

Qy   4602 GTGTGTGCGGGGTTTCAGCAATAGTTAAAGGCCAGCCGTAGGGGCCGTGGGA 4661
          |||||||||
Db   1355 CGACCGTGCCTCAAGACGCCGGTGTTCGCGCTCGCAGCGGGCGTAGGACTGGCAGGGGTA 1414

Qy   4662 GAGCTGGCATATACTATGTAGAGCGGAGTTGTACCCCTTCGGGTATGGTTCCTGAAT 4721
          ||||||||
Db   1415 GGAGAGGCCATCTACAAGGTTTGTGACTCCAGGAGAACGGCCTTCGGGCCATGTTCGATTCCT 1474

Qy   4722 GCAACATTGTTGAAGCCTTCGACGGCAGCCAAGGCATGTATGGTTTGTCATCAACAGAAG 4781
          ||||||
Db   1475 CGGTCCTATGTGAGTGCTATGACCGGGTATGCTTGGTATGAGCTCACCCGCCTGGTGGTTC 1534

Qy   4782 CTCAAACTATTCTTGGACACCTATCTGCGATAGGAGCAAATT 4841
          ||||||||
Db   1535 CCACGGTCAGATTGCCGGGTTACCTAAACACACAGGGGTTGCCGTCTCGTCAGGACCATC 1594

Qy   4842 TGGACGAGTGGGGCTGATCTCTTTTCTATGGTCAACCCC 4879
          ||||
Db   1595 TGGAGTTCGGGGAGGGGGTCTTCTTCACGGGGCCTCACCCAC 1632
```

RESULT 14
US-10-328-206-1
; Sequence 1, Application US/10328206
; Publication No. US20030162169A1
; GENERAL INFORMATION:
; APPLICANT: Pellerin, C. and Lamarre, D.
; TITLE OF INVENTION: Surrogate cell-based system and method for assaying the
; TITLE OF INVENTION: activity of Hepatitis C virus NS3 protease
; FILE REFERENCE: 13/075-2-C1
; CURRENT APPLICATION NUMBER: US/10/328,206
; CURRENT FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: US 60/132,360
; PRIOR FILING DATE: 1999-05-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 5211
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: partial cDNA
; OTHER INFORMATION: sequence of HCV (NS3-5B') fused to tTA
; OTHER INFORMATION: transactivator
US-10-328-206-1

Query Match   2.8%;   Score 262.4; DB 17;   Length 5211;
Best Local Similarity 49.6%;   Pred. No. 2.8e-65;
Matches 733;   Conservative   0;   Mismatches 736;   Indels   9;   Gaps   2;

Qy 3411 ACAACGTGTTGTATACTGCTCACCATGGCAGCAGCAAGGGGGCGCCGGTTGGCTCATCCCCACAG 3470
Db   155 ACGGCGGTGTGTTGGACTGTCTTCCATGGCGCCGGCTCAAAGACCTTGGCCGCCCCAAAG 214

Qy 3471 GCTCTATACACCCAATAACCGTTGACGCGGCTAATGACCAGCATCTATCAACCACCAT 3530
Db   215 GCCCGATCACCCCAGATGTACACTAATGTGACCAGGACCTGTCGCGGTGGCAGGCGCCCC 274

Qy 3531 GTGGAGCTGGGGTCCCTTACTCGGTCTCTTCGGTGGGGGAGGACCAAGGGTATCTGGTAACAC 3590
Db   275 CTGGGGCGCGCCGCTTCCATGACACCACCATGCGACCTGCGGCAGCTCGGCGACCTCTATTTGGTCACGA 334

Qy 3591 GACTGGGGGTCATTGGTTGAGTGAACAAATCCGATGACCCTTATTGGTGTGTGTGGCGGGGG 3650
Db   335 GACATGCCGACGTCAATTCGGGCGCGCCGGGCGGGGACAGTAGGGGGAGCCTGCTCTCCC 394

Qy 3651 CCCTTCCATGCGTCGTTGCCAAGGGTTCTTCAGGTGCCCCCGATTCGTGTCGTCCTCCGGGC 3710
Db   395 CCAGGCCTGTCTCCTACTTGAAGGGCTCTTCGGGTGCCCACTGCTCTGCCCCTTCGGGGC 454

Qy 3711 ATGTTATTGGATGTTCACCCGTGCTAGAAAATTCTGGCGGTTCAGTCAGTCAGATTAGGG 3770
Db   455 ACGCTGTGGCCATCTTCCGGGCTGCTGTGTGCACCCGGGGGTTGCAAAAGCGGTGACT 514

Qy 3771 TTAGGCCGTTGGTGTGTGGTCTGATCATGCCAACCATCCCCAGTACACAGCACATGGCCACTCTTGATA 3830
Db   515 TCATACCTGTTGAGTGTATGAAACTACTATGCGGTCTCCGGTCTTCCACGAGCAACTCAT 574

Qy 3831 CAAAACTACTGTGCCTAACGAGTATTCAGTGCAAATTTTAATTGCCCCACTGGCCGACGCG 3890
Db   575 CCCCCCAGCCGTACCGAGACATTCCAAGTGCCCCATCTACACGGTCCTACTGGCCAGCG 634

Qy 3891 GCAAGTCAACCAAATTACCACTTCCTTCATGCAGGAGAAGTATGAGGTCTTGGTCCTAA 3950
Db   635 GCAAGAGCACTCGTCAAAGTGCCGGCTTATGCGAGCAGCCAAGGGTACCTTGTCTTGA 694

Qy 3951 ATCCCAGTGTCGGCTACAACAGCATCAATGCCAAAGTACATGCCAAGTACGACGTACGGCGGTGA 4010
Db   695 ACCCGGTCCGTTGCCGGCCACCCTTAGGTTTGGGGCGTATATGTCAAAGGCACGTGGCACCG 754

Qy 4011 ATCCAAATTGCTATTTTAATGCAAAATGTACCAACACAGGGCTTCACTTACGTACAGCA 4070
Db   755 ACCCTAACATCAGAACTGAACTGGGGTAAGGACCATCACCACCAGCGCCCCCATCACGTACTCCA 814

Qy 4071 CATATGGCATGTACCTGACCGGAGCATGTTCCC------GGAACTATGATGTAATCATT 4124
Db   815 CCTATGGCAAGTTCCTTGCCGACGGCGTCTGTTTCTGTCTGGGGGGCGCCTTATGCACATCATAATGT 874

Qy 4125 GTGACGAATGGCCATGCTACCGATGCAACCACCCGTGTTGGCATTGGAAAGGTCTAACCG 4184
Db   875 GTGACGAGTGCCCACTCAACTGACTACCATCTGGGCATCGGCACCAGTGCCTCTGGACC 934

Qy 4185 AAGCTCCATCCAAAAATGTTAGGCTAGCTAGTGGTTCTTGCCACGGCTACCCCCCCTGGAGTAA 4244
Db   935 AAGCGGAGACGCGCTGGAGCCGGGCCTTGTCGTCGGCCGCACCGCTCCGGGATCGG 994

Qy 4245 TCCCTACCACCACATGCCAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCT 4304
Db   995 TCACCGTGCCCACCCCAATATCGAGGAGGTGGCCCTGTCCAACACTGGAGAGATCCCCT 1054

Qy 4305 TTCATGGAAAAGATTAAGGAGGAAAAATCTGAAGAAGGGGAGACACCTTATCTTGAGG 4364
Db 1055 TCTATGGCAAAGCCATCCCGAGTCAGTAGTCTGAAAGGGAATAACAGCTG 1114

Qy 4365 CTACCAAAAACACTGTGATGAGCTTGCTAACGAGTTAGCTGAAAGGGAATAACAGCTG 4424
Db 1115 ATTCCAAAAAGAGTGTGATGAGCTCCGCCAAAGCTGTCGGGCCTCAACGCCG 1174

Qy 4425 TCTCTTACTATAGGGGATGTGACATCTCAAAAAATCC---CTGAGGGGGACTGTGTAGTAG 4481
Db 1175 TAGCGTATTACCGGCGGCCTTGATGTGTCCGTCATACCGACCAGCGGAGACGTCGTTGTTG 1234

Qy 4482 TTGCCACTGATGCTTGTGTACAGGGTACACTTGTGACTTTGATTCCGTGTATGACTGCA 4541
Db 1235 TGGCAACAGACAGCGCCTTAATGACCGGCCTTTACCGGGCGACTTTCAGTCAGTGACTCA 1294

Qy 4542 GCCTCATGTGCTGTAGAAGGCACATGCCATGTGACCTTGACCCTACTTTCACCATGGGTGTTC 4601
Db 1295 ATACGTGTTTACCGACAGTCGATTTCAGCTTGGACCCCACCTTCACCATTGAGACGA 1354

Qy 4602 GTGTGTGCGGGGTTTCAGCAATAGTTAAAGGCCAGGCCGCACAGGCCGTGGGA 4661
Db 1355 CGACCGTGCCTCAAGACGCGGTGTCGCGGCTCGGCCAGGGCGGGTAGGACTCGGCAGGGGTA 1414

Qy 4662 GAGCTGGCATATACTACATGGGTGGCGGCCGGAGTTGTACCCCTTCGGGTATGGTTCCTGAAT 4721
Db 1415 GGAGAGGCCATCTTCACAGGTTTGTGACTCCAGGAGGAACCGGCCTTCGGGGCATGTTCGATTCCT 1474

Qy 4722 GCAACATTGTTGAAGCCTTCGACGCAGCCAAGGCATGTGTATGGTTTGTCATCAACAGAAG 4781
Db 1475 CGGTCCTATGTGAGTGCTTGTATGACGGGCGTGTCTTGGTATGAGCTCACCACCGCCTGAGA 1534

Qy 4782 CTCAAACTATTCTGACACCTATCGCACCAACCTGGGTTACCTGCGGATAGGAGCAAATT 4841
Db 1535 CCACGGTCAGATTGCGGGCTTACCTAAACACACCAGGTTGCCCGTCTGTCAGGACCATC 1594

Qy 4842 TGGACGAGTGGGCTGATCTCTTTTCTATGGTCAACCCC 4879
Db 1595 TGGAGTTCTGGGAGGGGCGTCTTCACGGGGCCTCACCCAC 1632

RESULT 15
US-11-140-379-1
; Sequence 1, Application US/11140379
; Publication No. US20050227226A1
; GENERAL INFORMATION:
; APPLICANT: Pellerin, C. and Lamarre, D.
; TITLE OF INVENTION: Surrogate cell-based system and method for assaying the
; TITLE OF INVENTION: activity of Hepatitis C virus NS3 protease
; FILE REFERENCE: 13/075-1-D1
; CURRENT APPLICATION NUMBER: US/11/140,379
; CURRENT FILING DATE: 2005-05-27
; PRIOR APPLICATION NUMBER: US 60/132,360
; PRIOR FILING DATE: 1999-05-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1

```
; SEQ ID NO 1
; LENGTH: 5211
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: partial cDNA
; OTHER INFORMATION: sequence of HCV (NS3-5B') fused to tTA
; OTHER INFORMATION: transactivator
US-11-140-379-1

Query Match        2.8%;   Score 262.4;   DB 26;   Length 5211;
Best Local Similarity  49.6%;   Pred. No. 2.8e-65;
Matches  733;   Conservative   0;   Mismatches  736;   Indels   9;   Gaps    2;

Qy   3411 ACAACGTGTGTATACTGCTCACCATGGCAGCAGCAAGGGCGCCGGTTGGCTCATCCCACAG 3470
           ||||| ||| || |||||||    || |||  | |     ||  | |    ||| ||| ||
Db    155 ACGGCGTGTGTTGGACTGTCTTCCATGGCGCCGGCTCAAAGACCTTGGCCGCGCCCAAAG  214

Qy   3471 GCTCTATACACCCAATAACCGTTGACGCGGCTAATGACCAGGACATCTATCAACCACCAT 3530
           | |||  | |  |  |   |||| |  | |   | | |    ||||  |  |  | || ||
Db    215 GCCCGATCACCCCAGATGTACACTAATGTGACCAGGACCTCGTCGGCTGGCCACGGCGCCC  274

Qy   3531 GTGGAGCTGGGTCCCTTACTCGGTGCTCTTGCGTGGGAGACCAAGGGTATCTGGTAACAC 3590
           | |||||  | |||  |  |  | | |   |||  || ||| |||  ||   ||| ||| |
Db    275 CTGGGGCGCGCTCCATGACACCTGCGGCAGCTGCAGTCGGACCTCTATTTGGTCACGA  334

Qy   3591 GACTGGGGTCATTGGTTGAGGTCAACAAATCCGATGACCCTTATTGGTGTGTGTGCGGGG 3650
           ||   |  ||| | |||| |||||  |  | ||||  |  | |     |   |   |  ||
Db    335 GACATGCCGACGTCATTCCGGTGCGGCCGGCGGGGGGACAGTAGGGGGAGCCTTGCTCTCCC  394

Qy   3651 CCCTTCCATGCGTGTTGCCAAGGTTCTTCAGTGCCCGATTCTGTCTCCTCCGGGC 3710
           ||  |   |  | |   |  ||||  |  | |   | |  |  || | |  | |||
Db    395 CCAGGCCGTGTCTCCTACTTGAAGGGCTCTTCGGGGTGGCGCCCTTCGGGGC  454

Qy   3711 ATGTTATTGGGATGTTCACCGTGCTAGAAATTCTGGCGGTTCAGTCAGTCAGATTAGGG 3770
           | | |  | | ||  || |  | |||| |    || ||  ||  |  |  | | || |||
Db    455 AGCTGTGGGACATCTTCCGGGCGTCTGTGTGCAACGGTTGCACCCGGGGGTTGCAAAAGCGGTGGACT  514

Qy   3771 TTAGGCCGTTGGTGTGTGTGCTGCTGGATACCATCCCCAGTACACAGCACATGCCACTCTTGATA 3830
           |  | |||| |    | | |||  |  || ||||| |   |     | | | ||  | ||| |  |
Db    515 TCATACCTGTTGAGTGTTATGGAAACTACTAATGCGGTCTTCCGGTCTCCGGTCTTCACAGACAAACTCAT  574

Qy   3831 CAAAACCTACTGCGTGCCTAACGAGTGATTCAGTGCAAATTTAAGTGCCCCACGCACGCG 3890
           ||     ||  ||  || |||  || |  |   | | | ||  ||| |   | | | |  |
Db    575 CCCCCCAGCCGTACGACCGCAGACATTCCAAGTGGCCCATCTCACACGGCTCCTACTGGCAGCG  634

Qy   3891 GCAAGTCAACCAAATTACCACTTTCTTTACATGCGAGGAAGTATGAGGTGTTGGTTCTTAA 3950
           ||    || |   |   |  || |||| |  | |  || || | ||| |  || || ||| |
Db    635 GCAAGAGCACTAAAGTGCCGGTGCTTATGCAGCCCAAGGTACTTGTTCTTGA  694

Qy   3951 ATCCCAGTGCTGGCTACACAACAGCATCAATGCCAAAGTACATGCGACGCTACGGCGTGA 4010
           |  | |  |||||   | |   |  |||  |||  ||| |||   |   |  |||| |  |
Db    695 ACCCGTCCGTTGCCGCCACCTTAGGTTTGGGCGTATATGTCAAAGGCACGTGGCCACCG  754

Qy   4011 ATCCAAATTGCTATTTTAATGCAAAATGTACCAACACAGGGGCTTCACTTACGTACAGCA 4070
           |  |||      | | |  | |    ||  | ||| ||| |  ||   | ||  |  | ||
Db    755 ACCCTAACATCAGAACTGGGGTAAGGACCATCACCAGGCCGCCCCCATCACGCGTACTCCA  814

Qy   4071 CATATGGCATGTACCTGACCGGAGCATGTTCCC------GGAACTATGATGTAATCATTTT 4124
           | ||||| |  | | ||  | |  |  | ||||      ||    || || || | |  |
Db    815 CCTATGGCAAGTTCCTTGCGACGGTGGTTGCTCGGGGGGCGCTCTGGCATTGGGAAGGTCCTAACCG  874

Qy   4125 GTGACGAATGCCATGTACCGGATGTCAACCACCGTGTTGGGCATTGGAAAGGTTCCTAACCG 4184
           || | |   || | ||  |  |  |  |   ||  |  || |   ||| ||  || ||| ||
Db    875 GTGACGAGTGCCACTCACGACTCGACTAGCATACCATCTTGGGCTCGTCGTCGCTCCCACCGGACC  934

Qy   4185 AAGCTCCATCCAAAAATGTTAGGCTAGTGGTTCTTCTTGCCGACCCCCCTGGAGTAA 4244
           || |   || ||   |   |   | ||| | ||      | |  || |  |  |  ||
Db    935 AAGCGGAGACGGCTGGAGCGCGGCTTGCCGTCTCGTGTCGTGCCCACCGTACGCCTACGGCTCCGGGATCGG  994

Qy   4245 TCCCTACACCACCATGCCAACATAACTGAGATTCAATTAACCGATGAAGGCACTATCCCCT 4304
           |||||  |  || || ||     ||| ||  ||  ||  | | || | ||  | || || |
Db    995 TCACCGiCCACAiCCCCATATCGAGGAGGTGGCCTGTGCCAACACTGAGACACTGGAGAGATCCCCT 1054
```

```
Qy   4305 TTCATGGAAAAAAGATTAAGGAGGAAAAATCTGAAGAAGAAATCTGAAGGGACCACCTTATCTTTGAGG 4364
           |||| ||||  ||||  ||||  |  || |   |  || |  ||| |||  |   || | || | |||
Db   1055 TCTATGGCAAAGCCATCCCATCGAGGTCATCAAGGGCATCAAGGGGAAGGCATCTCATTTTCTGCC 1114

Qy   4365 CTACCAAAAAACTGTGATGAGCTTGCTCTAACGAGTTAGCTCGAAAGGAATAACAGCTG 4424
           |  |||| |||| || ||  |||| |   |   ||  || |  |     | |  | | |
Db   1115 ATTCCAAAAAGAGAGTGTGATGAGCTCGCTGATTTCAGTTGGGGCCTCAAGCTGTCGGGGCCTCAACGCCG 1174

Qy   4425 TCTCTTACTATAGGGATGTGACATCTCAAAAATCC---CTGAGGGCGACTGTGTAGTAG 4481
           |  | | | | | ||  | |   |    |  | |      |  |||  |  | ||| |  |
Db   1175 TAGCGTATTACCGGGCCTTGATGTGTCCGTCATACCGGAGCCGGGCTGGCAGACGTCGTTGTTG 1234

Qy   4482 TTGCCACTGATGCCTTGTGTACAGGGTACACTGGTGACTTTGATTCCGTGTATGACTGCA 4541
           |||| || |  || ||   |  || |||| || | |   |||  || | ||| ||||||||
Db   1235 TGGCAACAGACGCGCCTTAATGACCGGCTTTACCGCGACTTTGATTCAGTGATCGACTGCA 1294

Qy   4542 GCCTCATGTAGAAGGCACATGCCATGTTGACCTTGACCTTGACCCTACTTCACCATGGGTGTTC 4601
           | | || |  ||   ||| |  |  ||| |  ||   |||  | ||  | |||  | ||| | |
Db   1295 ATACGTGTTACCCAGACAGTGCGATTTCAGCTTGACCCCACCTTCACCATTGAGACGA 1354

Qy   4602 GTGTGTGCGGGGTTTCAGCAAGTTAAGGCCGGCCACAGGCCGTGGGA 4661
           | | | | | |  |   | | || |  | | | || | | |  | |||
Db   1355 CGACCGTGCCTCAAGACGCCGTGTCGCCGGCTCGCCGCGGCGGGGTAGGACTGGCAGGGGGTA 1414

Qy   4662 GAGCTGGCATATACTACTATGTAACGGGAGTTGTACCCCTTCGGGTTCCTGAAT 4721
           |||  ||||| | |   | | |  |||  |||| |  | || ||| || |  |
Db   1415 GGAGAGGCATCTACAGGTTTGTGACTCCAGGAGAACGGCCTTCGGGGCATGTTCGGATTCCT 1474

Qy   4722 GCAACATTGTTGAAGCCTCGACGCAGCCAAGGCATGATGGTTGTCATCAACAGAAG 4781
           | | ||||| | |||  |  |   |  | |||  | | ||| |||  | || ||
Db   1475 CGGTCCCATGTGAGTGCTTATGACGGGCTGTCGTGTTGGTTATGAGCTCACGCCGCTGAGA 1534

Qy   4782 CTCAAACTATTCTGGACACCTATCGCACCCAACCTGGGTTACCTGCGGATAGGAGCAAATT 4841
           | |||| || | ||| ||| |  ||||| |  ||| || || | || | |  ||| |   |
Db   1535 CCACGGTCAGATTGCGGGTTACCTACCAGCAGCCAGGGTGTCCCGTCTGTGTCAGGACCATC 1594

Qy   4842 TGGACGAGTGGGCTGATCTCTTTTCTATGGTCAACCCC 4879
           |||| ||||| | || ||| || | | | || | |||
Db   1595 TGGAGTTCGGGGACGGGCCGTCTTCACGGGGCCTCACCCAC 1632
```

Search completed: October 31, 2005, 00:41:00
Job time : 4391 secs